Master thesis

Recurrent origins and adaptive advantage of polyploid *Parnassia palustris*

Hanna Hagen Bjørgaas

December 1st 2011

CEES

Centre for Ecological and Evolutionary Synthesis

Department of Biology, University of Oslo

Forord

Det er mange som fortjener en stor takk i dette forordet. Jeg har banket på mange dører, sendt utallige mailer – og fått like mange positive og hjelpsomme svar, både under feltarbeidet, på lab, og i analyse- og skriveprosessen. Viktigst er mine fine veiledere Anne, Aud og Marte, som jeg tror må ha hatt de åpneste veilederdørene på Nedre Blindern. Takk for denne fine oppgava (selv om modellarten har nykker som en primadonna), for tillit og boller, og takk for alle lånene av herr Trøst!

Takk til Leif for å ha vært tålmodig assistent/sykepleier under feltarbeidet. Thanks to Marlene, Magnus, Sachiko, Gunnløg, Einar, Mette, Bjørgaasene and everybody else who has sent material, helped me in the field or at the lab. Nanna, Emelita og Cecilie skal ha takk for å være til stor hjelp på lab. En stor takk rettes til Kyrre, Trond, Lars og Geir, som har gjort statistikk forståelig og anvendelig for meg. Takker også Marit og Anna, samt Idunn, Siri, Babsi og resten av andre etasje for vin og fliring.

Contents

Abstract

Background and aims: Polyploidy is a major driving force behind angiosperm evolution. The distribution of polyploid taxa has lead previous authors to suggest a connection between genome duplication and major climatic events, like glaciations. In this study, European di- and tetraploid *Parnassia palustris* were used to explore how the genetic structure and distribution of the cytotypes relates to the glacial history of the study area, to evaluate the taxonomic treatment of the cytotypes, to search for evidence for interploidal gene flow and/or recurrent formation of the tetraploid, and to explore differences in ecological flexibility between cytotypes.

Material and methods: Flow cytometry and AFLP were used to estimate ploidal levels and genetic structure. Growth experiments were performed in search for effects of ploidy on growth and survival during various day lengths and temperature treatments, employing Generalised Linear Models (GLMs) and Survival analyses.

Results: Tetraploid populations were recorded from the Alps for the first time, and triploids were detected in areas where diploid and tetraploid populations are sympatric. Ploidal levels explained little of the genetic variation, and di- and tetraploid populations from the same area were found to be genetically similar. High levels of genetic diversity were found in Fennoscandian populations. According to GLMs and survival analyses, survival was the only out of several fitness measures that was significantly affected by ploidy. High temperatures resulted in lower tetraploid survival.

Conclusions: Taxonomic rank for the tetraploid cytotype could not be supported. The distribution of cytotypes (with tetraploids mainly distributed in previously glaciated areas), the admixed genetic structure and higher levels of diversity found in Fennoscandian populations of *P. palustris* all suggest a scenario with multiple immigration lineages from at least two refugia that met in a contact zone in Fennoscandia. Gene flow between the cytotypes and/or recurrent origins of the tetraploid are likely to have contributed to the high levels of genetic diversity. However, no differences in ecological flexibility could be attributed to polyploidy *per se* based on the growt experiments.

Introduction

Polyploidy is a stimulus for flowering plant evolution, and has played a part in the evolutionary history of nearly, if not all angiosperm lineages (Levin 2002; Soltis, Albert et al. 2009; Soltis, Buggs et al. 2010). Genome duplication is a widely accepted speciation mechanism, the most important single mechanism causing sympatric speciation in land plants (Adams and Wendel 2005; Hendry 2009) and is associated with remarkable rapid evolution and variation in land plants (Pires, Zhao et al. 2004). Polyploidisation events are suggested to be linked to climatic changes, and conquering of new habitats and niches (Favarger 1967; Fawcett, Maere et al. 2009; te Beest, Le Roux et al. 2011). Understanding how climatic processes affect polyploidisation events, and the consequences of such events, can provide insight into plant speciation and diversity at a larger scale (Parisod and Besnard 2007). However, important aspects of mechanisms behind and consequences of polyploidisation remain poorly understood (Ramsey and Schemske 1998; Soltis, Buggs et al. 2010).

Typically, polyploids are categorized as allopolyploids or autopolyploids (Stebbins 1947). However, also the definitions of these labels have generated controversy for decades, and are commonly considered as extremes on a scale with many intermediates (Grant 1981; Soltis, Soltis et al. 2003; Soltis, Buggs et al. 2010). Following a taxonomic definition, a polyploid arising within a single species will be defined as autopolyploid, while a hybridisation between two different species, stabilized by a genome duplication, gives rise to an allopolyploid (e.g. Soltis, Soltis et al. 2007). Following a cytological definition, chromosomal behaviour during meiosis plays the important part: In principle, an autopolyploid will exhibit multivalent pairing, while an allopolyploid will exhibit bivalent pairing (Darlington 1932; Ramsey and Schemske 1998). The taxonomical treatment of autopolyploids depends on which species concept one chooses to use (see e.g. Wheeler and Meier 2000; Soltis, Soltis et al. 2007) and many autopolyploids are treated merely as cytotypic variations of their parental species (Soltis, Soltis et al. 2007).

Polyploidisation has often been regarded as a form for instant speciation, as reproductive isolation is expected to occur immediately after a polyploidisation event because of the reduced fertility of a triploid hybrid (Husband 2004; Hendry 2009). However, even a

partly fertile triploid could form a "bridge" between the polyploid and its progenitors, allowing gene flow between the cytotypes (Husband 2004). Recurrent formation of polyploids is considered the rule rather than the exception (Soltis and Soltis 1999; Levin 2002; Soltis, Soltis et al. 2003). Thus, both recurrent origins of the tetraploid cytotype and hybridisation between unreduced diploid gametes and reduced tetraploid gametes could lead to gene flow between ploidal levels (Stebbins 1971; Slotte, Huang et al. 2008; Soltis, Buggs et al. 2010; Jørgensen, Ehrich et al. 2011). However, the rate and role of these processes remain poorly understood in natural plant populations (Soltis, Buggs et al. 2010).

Since examinations of the angiosperm *Oenothera lamarckiana* Ser. mut. *gigas* (Onagraceae) revealed chromosome doubling in the early 20th century (Lutz 1907; Gates 1909; Soltis, Buggs et al. 2010), botanists have recognized that the proportion of polyploid taxa often increases towards higher latitudes (Müntzing 1936; Flovik 1940; Löve and Löve 1943) and that many polyploid taxa occupy previously glaciated territories (Favarger 1967; Abbott and Brochmann 2003). Recently evolved high-ploid species complexes are overrepresented in the previously glaciated Arctic area (Soltis and Soltis 1999; Abbott and Brochmann 2003) when compared to the largely unglaciated Beringian area (Brochmann, Brysting et al. 2004).

Polyploid plants often exhibit traits that differ from those of their progenitors (Levin 2002; Ramsey and Schemske 2002), and various explanations for such geographical and ecological differences between cytotypes have been posed. Early authors suggested that because of greater genetic variability compared to their diploid progenitors, polyploids were more tolerant to harsh and fluctuating environments (e.g. Hagerup 1931; Müntzing 1936; Löve and Löve 1943; Manton 1950; Johnson and Packer 1965). Flovik stated that "*an increased chromosome number increases the adaptability to extreme habitats, like arctic and alpine conditions*" (page 439; in Flovik 1940). More recent authors have suggested that polyploids may have an evolutionary advantage only if the conditions are drastically changing, e.g. during the Cretaceous-Tertiary extinction event at about 65 mya (Fawcett, Maere et al. 2009). Perhaps one of the most commonly accepted hypotheses explaining the distribution differences between cytotypes concerns secondary contact between lineages: Stebbins (1984) proposed that hybridisation between populations that

had diverged in different refugia during cold periods, and then met in a contact zone, gave rise to new gene combinations. Stabilized by polyploidisation, some of these new hybrids could harbour favourable gene combinations, and therefore be better adapted to the new and deglaciated habitat than their diploid progenitors. This view was supported by Soltis and Soltis (2000), who suggested that the higher levels of heterozygosity found in polyploids could cause an increased ecological flexibility compared to their diploid progenitors. Intragenomic rearrangements and altered gene regulation following genome doubling may contribute to novel expression phenotypes and evolutionary flexibility in both allo- and autopolyploids (Levin 1983; Levin 2002). Other authors have argued for non-adaptive reasons behind differences in distribution patterns of cytotypes. Such explanations include "self cleaning" of mixed populations due to low triploid fitness, in addition to merely historical reasons (Dijk and Schotsman 1997). Most studies have examined polyploid systems that diverged thousands of years ago, and critics claim that it is impossible to separate the effect of polyploidisation *per se* from that of genetic evolution in these systems (Ramsey 2011). Thus, inferring an adaptive advantage of polyploids under certain conditions is not uncontroversial.

Because of diverging distribution patterns, in addition to the observed physiological and ecological differences between cytotypes, it has been suggested that polyploidy mediates ecological differentiation and speciation, also in the case of autopolyploids (Levin 2002; Ramsey 2011). However, most studies concerning possible adaptive differences between cytotypes have so far been observational (Soltis, Buggs et al. 2010) and little is known about the competitive status of polyploids (Levin 2002). The few experiments that have been conducted to explore differences in performance between cytotypes have given ambiguous results (Stebbins 1949; Stebbins 1985; Bretagnolle and Lumaret 1995; Baack and Stanton 2005; Münzbergová 2006; but see Ramsey 2011).

Genetic structure, meaning quantity and distribution of genetic variation within and among populations (Excoffier 2007), in a species distribution range is often a result of a combination between historical events and contemporary processes. Researchers agree that most of Fennoscandia was covered by ice from Middle Weichselian to the deglaciation at approximately 11,500 years ago, with a last glacial maximum (LGM) occurring at about 21,000 years before present (Lokrantz and Sohlenius 2006).

Postglacial immigration of the North Atlantic region from multiple refugia after the LGM is shown to have effects on a species' distribution of such variation (Schönswetter and Tribsch 2005; Alsos, Eidesen et al. 2007; Meirmans, Goudet et al. 2011). Investigating the genetic structure of a population can therefore shed light over the past and present processes affecting the population's present distribution range.

Parnassia palustris L. (Linnaeus 1753:273) is the only widespread species in the genus *Parnassia*, which comprises about 70 species (Hultgård 1987; Wu, Wang et al. 2005). The genus is now commonly placed in Parnassiaceae (Elven 2007 onwards) although the family affiliation has been discussed (Hultgård 1987; Li-Bing and Simmons 2006). Flower anatomy and crossing experiments strongly suggest that the species is mainly outcrossing (Hultgård 1987). Two cytotypes of *P. palustris* are commonly found: $2n = 18$ and $2n = 36$ (Erlandsson 1942; in Hultgård 1987). Several studies suggest that the tetraploid cytotype is a result of more than one autopolyploidisation event (Hultgård 1987; Borgen and Hultgård 2003). The two cytotypes of *P. palustris* have a somewhat different distribution in Europe. The diploids are predominant in Southern Scandinavia and southwards to Central and Southern European mountains, while the tetraploids are found mainly in Northern Fennoscandia. The two cytotypes are occurring in sympatry in a zone across Southern Norway and Sweden (Hultgård 1987), mainly corresponding to the glacial boundary approximately 11,500 years ago (Brochmann, Brysting et al. 2004). This distribution pattern has lead previous authors to suggest an advantage of the tetraploid cytotype in colonizing the previously glaciated area (Hultgård 1987; Borgen and Hultgård 2003).

Because of the divergent but overlapping distribution patterns of the two cytotypes and the association of the autotetraploid cytotype with previously glaciated areas in Fennoscandia, *P. palustris* is a suitable model species for examining the ecological consequences and the evolutionary history of polyploidisation in a postglacial landscape. Using AFLPs, flow cytometry data and *P. palustris* samples from throughout Europe, I will address the following questions:

- Do genetic analyses support species rank for the European tetraploid cytotype?
- Is the distribution of cytotypes consistent with previous studies, with the tetraploid cytotype prevailing in previously glaciated areas?
- Do genetic analyses suggest multiple polyploid origins of and/or hybridisation between cytotypes*?*
- Did the Fennoscandian populations origin from one or more source regions?
- Do tetraploid plants of *P. palustris* display higher ecological flexibility than diploid plants?

Materials and methods

Plant materials

Seeds, plants and silica dried material of *Parnassia palustris* were collected from European, Alaskan and Japanese populations between August 2007 and October 2010 (Fig. 1, Table 1, Table S1 in Appendix). Leaf material from six to ten plants per population was collected and stored on silica gel. Seeds from up to ten plants from each population were collected and stored cool and dark. If no mature seeds were available, up to five live plants were collected and grown in a climate regulated growth room with artificial light with daylight quality (Powerstar HQI-BT 400 W/D, OSRAM, Munich, Germany) in addition to daylight, and 18 h photoperiod in the Phytotron (University of Oslo, UiO) until mature seeds could be collected. Vouchers from 30 of the populations were pressed in the field or in the Phytotron, and deposited at the Natural History Museum, University of Oslo (O).

In order to obtain fresh plant material for flow cytometry, silica dried material for genetic analyses, and young plants for growth experiments, the collected seeds were germinated. Several methods were tested, including three different temperatures (7 °C, 15 °C and 22 °C), two different light regimes (dark/light) and presence/absence of 10 µM gibberellin GA_3 solution. The following procedure was found to give the highest percentage of germinating seeds. Petri dishes (9 cm diameter) were prepared by adding 5 ml of distilled water and 2 ml of 10 µM gibberellin GA_3 solution to multiple layers of cellulose paper and one filter paper (Whatman no. 1, GE Healthcare, Buckinghamshire, UK), before adding the seeds. One filter paper was put on the top of the seeds and 2 ml distilled water was added. The dishes were closed and placed in a climate regulated and environmentally controlled growth room with daylight at 19±1 °C, relative humidity above 60%, and artificial 18h light with daylight quality (Powerstar HQI-BT 400 W/D, OSRAM), giving 210 µmol $m^{-2}s^{-1}$ photosynthetic photon flux density (PPFD) at plant height, in addition to natural low autumn daylight at 60 °N. After approximately 14 days, seedlings were transplanted into pots (11 C, OS Plastic A/S, Frederiksborg, Denmark) with soil (Herbia Plantejord, Nordic Garden AS, Stokke, Norway) mixed with perlite in a 5:1 proportion.

Table 1. Sampled populations of *Parnassia palustris.* English names are applied to localities and sub-localities when available; otherwise, local names are used. Abbreviations of collectors are written out in Table S2 in Appendix. Altitudes are given as meter above sea level (m asl). Ploidy is estimated by flow cytometry on fresh and silica dried material (see results). Positive longitude and latitude means north and east, respectively, whereas negative longitude means west. The number of individuals successfully analysed by AFLP and flow cytometry is recorded for each population. All vouchers are deposited at the Natural History Museum, University of Oslo (O). N/A – not available information.

Pop. nr	Year	Country	Locality	Sub-locality	Collectors	m asl	Longitude	Latitude	AFLP	Flow	Ploidy	Voucher
1	2009	Iceland	Akranes	Elinarhöfði	MHJ, AKB, CEP	3	64.583	-22.1161	3	9	4x	HHB - 1
2	2009	Iceland	Grændalur	Suðvesturland	MHJ, AKB, CEP	74	64.173	-21.4561	3	8	4x	HHB - 2
3	2008	Norway	Akershus	Nittedal	MU	380	60.024	10.8101	1	16	4x	HHB - 3
4	2008	Norway	Oslo	Nordmarka	MU	290	59.899	10.0969	3	27	4x	HHB - 4
5	2008	Norway	Vest-Agder	Lista	KH	2	58.059	6.7911	2	2	4x	HHB - 5
7	2008	Norway	Oslo	Lillomarka	MU, IH	189	59.97	10.8099	3	5	4x	HHB - 7
8	2009	Norway	Hordaland	Finse	MHJ et al.	1300	60.605	7.4942	2	13	4x	HHB - 8
9	2007	Norway	Oppland	Bøverkin	HHG	1000	61.66	8.15	1	1	4x	HHB - 9
10	2009	Slovakia	High Tatras	Javorna	MHJ	499	49.251	20.1546	5	10	2x, 4x*	HHB - 10
11	2009	Slovakia	Low Tatras	Pusté Pole	MHJ	565	48.886	20.2408	4	4	2x	HHB - 11
12	2009	Norway	Oppland	Synna	MHJ, HL	800	9.942	61.074	3	17	4x	HHB - 12
13	2009	Switzerland	Obwalden	Engelberg	HHB, AKB	1250	46.825	8.4319	1	9	2x	HHB - 13
14	2009	Switzerland	Montreaux	Les Rochers de Naye	HHB, AKB	1570	46.614	7.0008	1	10	2x	HHB - 14
15	2009	Norway	Finnmark	Vardø	HHB, LB	4	70.397	31.0737	3	3	4x	HHB - 15
16	2009	Norway	Finnmark	Durevuoppi	HHB, LB	3	70.478	27.9758	5	14	4x	
17	2009	Norway	Finnmark	Børselv	HHB, LB	50	70.382	25.7436	5	12	4x	HHB - 17
18	2009	Norway	Finnmark	Indre Veines	HHB, LB	5	70.431	25.2864	4	9	4x	HHB - 18
19	2009	Norway	Finnmark	Cape North	HHB, LB	12	71.204	25.8217	2	1	4x	
20	2009	Norway	Finnmark	Honningsvåg	HHB, LB	10	71.133	26.0353	6	4	4x	HHB - 20

Pop. nr	Year	Country	Locality	Sub-locality	Collectors	m asl	Longitude	Latitude	AFLP	Flow	Ploidy	Voucher
21	2009	Norway	Finnmark	Alta	HHB, LB	251	70.262	24.3089	1	1	4x	HHB - 21
22	2009	Norway	Troms	Kvænangen	HHB, LB	488	69.91	21.47	2	3	4x	
23	2009	Norway	Troms	Lyngseidet	HHB, LB	50	69.549	20.0964	5	1	4x	HHB - 24
26	2009	Finland	Lapplands län	Konkamaalven	HHB, LB	401	68.593	21.7872	3	21	4x	
27	2009	Finland	Lapplands län	Muonio	HHB, LB	225	67.74	23.91	4	4	4x	
28	2009	Sweden	Norrbottens län	Muddus	HHB, LB	183	66.855	20.1242	1	2	4x	HHB - 28
29	2009	Sweden	Västerbottens län	Vilhelmina	HHB, LB	361	64.44	17.0653	0	1	4x	
30	2009	Norway	Sør-Trøndelag	Rørosvidda	HHB, LB	838	62.641	12.0786	1	7	4x	HHB - 30
31	2009	Norway	Hedmark	Einunndalen	HHB, LB	1036	62.4	10.1597	0	7	4x	HHB - 31
32	2010	Netherlands	Zijpe	Zvanenwater	PK	2	52.824	6.0742	4	2	2x, 5x*	
36	2009	Norway	Vest-Agder	Einarsneset	KH	4	58.059	6.7911	0	1	4x	
37	2009	Alaska	Fairbanks	Sheep Creek	DM	135	64.838	-147.716	5	17	2x	
38	2009	Norway	Østfold	Butjern	HHB, MN	183	59.487	11.7232	2	10	2x	HHB - 38
39	2009	Russia	Karelia	Petrozavodsk	JJ	92	61.937	34.2194	0	5	2x	
41	2009	Slovakia	Belianske Tatras	N/A	AR	1370	49.251	20.2222	3	5	2x	
42	2009	Slovakia	Low Tatras	Krakova hol'a	AR	1480	48.998	19.6161	3	5	2x	HHB - 42
43	2009	Slovakia	Slovak Paradise	Verarska Tiesnava	AR, MR	790	48.931	20.2897	3	5	2x, 3x*	HHB - 43
44	2009	Russia	Murmansk	Kirovsk	TF	348	67.614	33.6719	4	17	4x	
46	2009	Austria	Kärnten	High Tauern	AT	2430	47.076	12.838	4	5	2x	
47	2010	Austria	Tirol	Lechtal Alps	GP	2157	47.143	10.2117	4	3	2x	
48	2010	Austria	Tirol	Lechtal Alps	GP	1496	47.101	10.2147	4	3	2x	
49	2010	Slovenia	Gorenjska	Julian Alps	GP, AT	1700	46.234	13.9347	4	4	4x	
50	2010	Austria	Carinthia	Karawanken	GP, AT	2100	46.502	14.4947	4	5	4x, 5x*	
51	2010	Russia	Kolesme	Belomorsk	LS	32	64.546	34.7728	6	6	4x	HHB - 51
52	2010	Norway	Møre og Romsdal	Kallandsvågen	MB	20	63.333	8.5833	5	4	4x	
53	2010	Norway	Nordland	Vega	MB	16	65.705	11.8667	2	3	4x	
54	2010	Norway	Nord-Trøndelag	Leka	MB	10	65.081	11.6122	3	3	4x	

Pop. nr	Year	Country	Locality	Sub-locality	Collectors	m asl	Longitude	Latitude	AFLP	Flow	Ploidy	Voucher
55	2010	Italy	Trentino-Alto	Ahrntal	AH	1700	46.988	11.9053	6	5	2x	
57	2010	Denmark	Vendsyssel	Biersted Mose	EFJ	12	63.184	12.0839	4	5	2x	
58	2010	Denmark	Nord-Jylland	Tversted	EFJ	10	57.591	10.1546	0	5	2x	HHB - 58
59	2010	Denmark	Mols	Knebel Vik	EFJ	4	56.199	10.4534	4	4	2x	
60	2010	Estonia	Tartu	Luunja	AH	64	58.345	26.8281	1	3	4x	
61	2010	Sweden	Jönkoping	Vaggeryd	GJ	197	57.492	14.1179	2	2	4x	
62	2010	Scotland	Argyll and Bute	Lochgilphead	WM	60	56.075	-5.4333	4	6	4x	
65	2010	Sweden	Jönkoping	Vaggeryd	GJ	254	57.589	13.9169	0	4	4x	
66	2010	Sweden	Skåne	Stängby mosse	HHB, MP	23	55.776	13.1528	4	12	2x, 3x*	HHB - 66
67	2010	Japan	Aichi prefecture	Seto	HHB, SN	389	35.198	137.1339	1	3	2x	
68	2010	Scotland	Argyll and Bute	Beinn Ime	BM	900	56.233	-4.8167	2	2	4x	
70	2010	Sweden	Halland	Hasslöv	HHB, MP	195	56.367	12.983	0	1	2x	HHB - 70

* Ind 10_10 is 4x. Ind 43_1 is 3x. Ind 50_5 is 5x. Ind 66_5 and 66_6 are 3x. None of these individuals are included in the AFLP analyses. Ind 34_2 is 5x and included in the AFLP analyses.

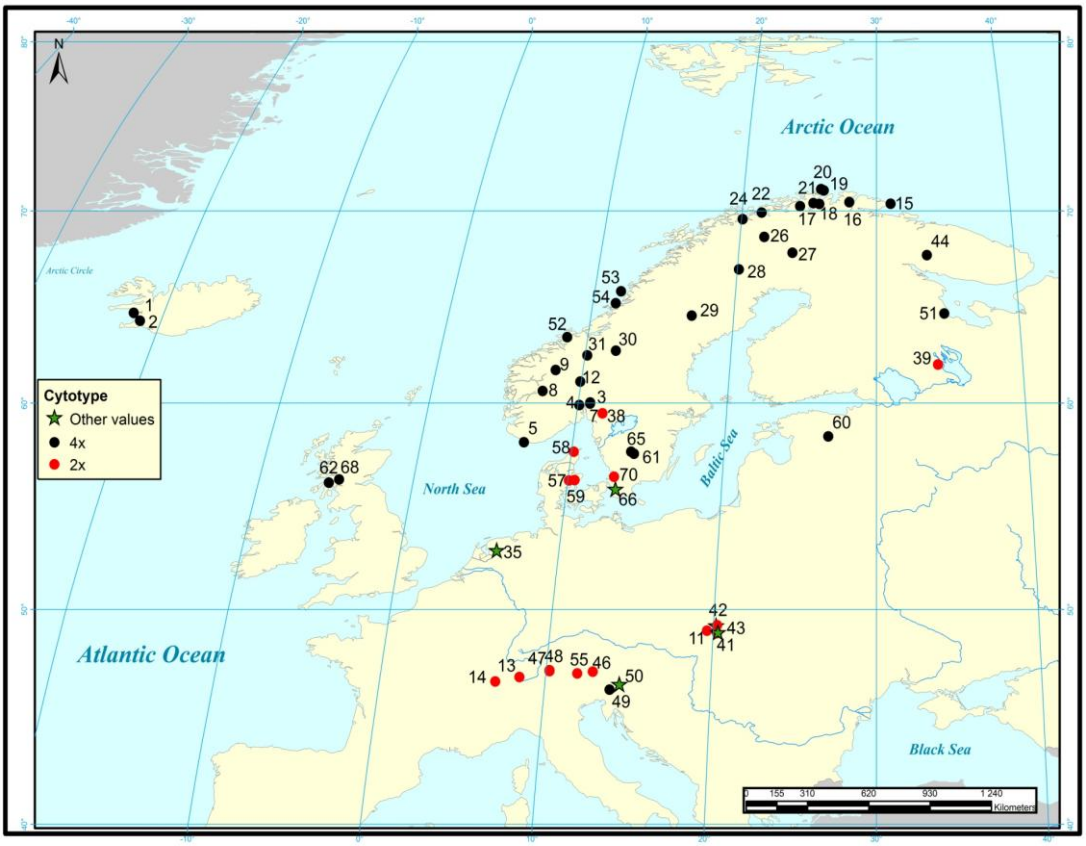

Figure 1. Sampled European populations of *Parnassia palustris*. Ploidy levels are based on flow cytometry of fresh and silica dried material (see results). Diploid populations are marked with red dots, tetraploid populations are marked with black dots. Populations, where more than one ploidal level were found, are marked with green stars.

Flow cytometry

DNA ratios from fresh material were obtained by G. Geenen, Plant Cytometry Services (Schjindel, The Netherlands) in November 2009 and April 2010, using *Ilex crenata* Thunb. 'Fastigiata' as internal standard and following the DAPI staining protocol as described in Jørgensen, Ehrich et al. (2006). Flow cytometry of silica dried material was performed by P. Travnicek, Laboratory of Flow Cytometry, Institute of Botany, Academy of Sciences of the Czech Republic (Průhonice, Czech Republic), using *Bellis perennis* L. as internal standard and following the protocol described in Bendiksby, Tribsch et al. (2011).

To check for reproducibility, 24 samples previously analysed from fresh material were also included in the analysis of silica dried material. In addition, six of the samples, representing both di- and tetraploids, were analysed from both fresh material and silica dried material. When ploidal level was not available for all plants or more than one ploidal level were found in a population, the dominating ploidal level of the examined plants was used as the overall ploidal level of populations (Table S1 in Appendix).

DNA extraction and AFLP

Amplified fragment length polymorphism (AFLP) is a marker system based on random cutting of total DNA. AFLPs are anonymous; fragments of the same size are not necessarily homologous (Mueller and Wolfenbarger 1999), and the markers are dominant; at a specific loci, the profile of a homozygous could be identical to that of a heterozygous (Kosman and Leonard 2005). It is therefore not unproblematic to infer genetic distances based on AFLP data, and AFLP data should be inferred as phenotypes, not as genotypes (Kosman and Leonard 2005). In spite of these shortcomings, the method is proved useful to provide intraspecific variation, and has gained popularity among plant phylogenetists and phylogeographers, e.g. (Stehlik, Schneller et al. 2002; Schönswetter, Suda et al. 2007; Westergaard, Jørgensen et al. 2010). Comparative studies have shown that AFLPs in large numbers are well suited to reveal true genetic diversity under certain evolutionary circumstances (Mariette, Corre et al. 2002).

Approximately 20 mg of silica dried material per plant were crushed in a mixer mill. Various methods for extracting DNA from the samples were tested, and the resulting DNA concentration was measured using a spectrophotometer (NanoDrop, ND-1000, Thermo Fisher Scientific, Wilmington, Massachusetts, USA). Protocols provided by Qiagen (Hilden, Germany), Omega Bio-Tek (Norcross, Georgia, USA) and Ziegenhagen, Guillemaut et al. (1993) did not show satisfactory yield and purity of DNA. For the final DNA extraction, a 2 x cetyl trimethyl ammonium bromide (CTAB) method (Murray and Thompson 1980) with the modifications described by Saghai-Maroof, Soliman et al. (1984) and Doyle, Doyle et al. (1987) was used. The DNA was eluted in 50 or 100 µl Tris-EDTA (TE) buffer and stored in a -20 °C freezer. Due to impurities in the DNA, an extra cleaning step using DNeasy Plant Mini kit (Qiagen) in accordance with steps 6 – 11 in the protocol was included with the following

modification: In the last step, the columns were heated to 65 °C and 30 µl of pre-heated TE buffer was added repeatedly twice to elute and concentrate the DNA. Agarose gel electrophoresis was used to visualise concentration and quality of each DNA sample, and adequate samples were analysed further.

In order to find primer combinations yielding a suitable amount of variation among the fragment length patterns, AFLP primer tests were conducted on duplicated individuals representing the geographic range of the sampled populations. Restriction-ligation (RL), preselective PCR and selective PCR were conducted in a DNA Engine Tetrad 2 Peltier Thermal Cycler (Bio-Rad Life Sciences, Hercules, California, USA) following procedures as described by Jørgensen, Elven et al. (2006) with the following exceptions: Each RL reaction was added 2 µl DNA template, giving a total reaction volume of 11 µl. To each preselective PCR reaction, 0.075 µl of AmpliTaq (Applied Biosystems, Foster City, California, USA) was added. The preselective elongation was 2 min at 71 ºC, and the selective elongation was 1 min at 72 ºC. The FAM-, NED-, VIC- and PET-labeled selective PCR products from each individual were mixed in a 3:2:2:2 ratio. To 2 µl of this mix, 13.4 µl Hi-Di formamide and 0.3 µl GeneScan Liz 500 Size Standard (both Applied Biosystems) were added. The samples were denatured for 5 min at 95 ºC before electrophoresis was performed using an ABI PRISM 3100 genetic analyser (Applied Biosystems).

Of the 18 primer combinations tested, the four combinations giving most appropriate number of reproducible peaks in the 50-500 bp range were run on the total set of samples. These were: FAM-labeled *Eco*RI-ACC – *Mse*I-CAT, PET-labeled *Eco*RI-AGA – *Mse*I-CTG, VIC-labeled *Eco*RI-AAG – *Mse*I-CTT and NED-labeled *Eco*RI-ACA – *Mse*I-CAC. In order to trace contaminations and ensure reproducibility, two individuals from extraction and 52 individiuals from RL were replicated. Individuals were also replicated between and within the 96-well plates. Unambiguous fragments in the size range 50-500 bp were visualised and scored as present (1) or absent (0) using the software GeneMapper 4.0 (Applied Biosystems). For each primer combination, the automatic scoring settings found to be most important (bin width, peak height treshold and minimum fragment size; Holland, Clarke et al. 2008) were adjusted to fit the profiles, and the resulting scoring was checked manually. Individuals with number of fragments deviating from the main range (30 - 50 fragments per profile) were carefully controlled and removed if they looked abnormal, as this could be a signal of

imperfect PCR. None of the negative controls showed patterns one could expect from contamination, and duplicates between and within runs were considered sufficiently similar. Populations and markers where the duplicated individuals showed dissimilarities were excluded from further analyses, giving a total of 144 markers from 155 individuals representing 53 populations. In four individuals, one of the primer combinations failed to amplify, resulting in missing data for these individuals (Table S1 in Appendix). The resulting data were exported as a presence/absence matrix.

Data analyses

The data set was examined for clones using the R script AFLPdat (Ehrich 2006) which was also used for most data format conversions, calculations of diversity and rarity values as described in detail later on. All analyses, except the Structure analyses, were performed both including and excluding populations from Alaska and Japan. Two Bayesian allocation approaches (BAPS and Structure, see below) were used to identify genetic groups in the dataset and to allocate individuals to these groups. Bayesian models calculate K groups of individuals with such composition that the individuals in each group have a genotype as similar as possible, and that the groups are as close to Hardy-Weinberg equilibrium as possible (Corander, Marttinen et al. 2008), resulting in a model describing an "optimal" number and composition of groups. The logarithmic probability of data given the model, and the posterior probability Pr(X|K) (equation 12 in Pritchard, Stephens et al. 2000) for the model, are calculated for each number and composition of groups (Evanno, Regnaut et al. 2005).

BAPS uses a stochastic estimation algorithm to infer the most likely number of groups (Corander and Marttinen 2006; Corander, Marttinen et al. 2008). Prior predictive distribution of the observed data, hereafter named the marginal likelihood, is calculated based on the posterior distributions over the space of putative clustering solutions. Individual level mixture analyses were performed by BAPS ver. 5.3, using no prior information about population origins.

In contrast to BAPS, Structure uses a Markov chain Monte Carlo (MCMC) algorithm to cluster individuals into groups and to find migrants and admixed individuals. The method is

developed for multilocus genotypic data (Pritchard, Stephens et al. 2000) but is also compatible with phenotypic AFLP data (Evanno, Regnaut et al. 2005). As Structure is developed for co-dominant markers, the dominant AFLP multilocus phenotypes were treated as diploid multilocus genotypes, adding a row of missing values for the unknown genotype (Eidesen, Alsos et al. 2007). The analyses were performed using Structure ver. 2.3.3 at the Bioportal, University of Oslo (freely accessible at www.bioportal.uio.no) with 10^6 iterations and a burn-in of 10^5 iterations. An admixture model, where fractions of the genome of an individual are allowed to descend from different groups (Pritchard, Stephens et al. 2000), was assumed. Linkage between loci was not taken into consideration. The number of groups K was varied between 1 and 9, in 10 independent runs for each number of K.

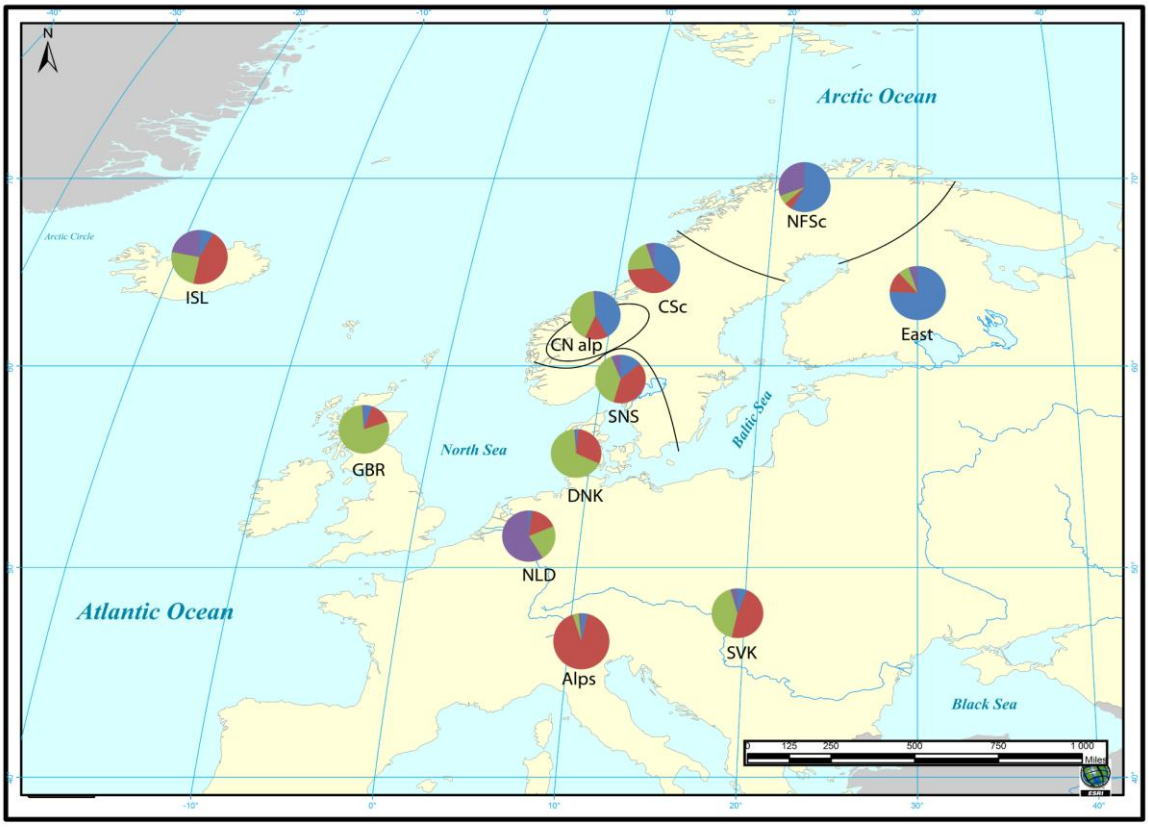

Figure 2. Sampled individuals of *Parnassia palustris* are assigned to geographical regions, here delimited by black lines, as defined in Table 3. Allocation analysis based on 144 AFLP markers clustered the European individuals into four clusters: Cluster 1(blue), Cluster 2 (red), Cluster 3 (green) and Cluster 4 (violet). The pie charts show the genetic structure of each region, i.e. the overall allocation to each of the four clusters for all individuals within each region.

The logarithmic probabilities of the observed data (Ln P(D)), similarity coefficients and ΔKs were calculated using StructureSum (Ehrich 2006), and were used to choose a number of K, as recommended in Pritchard, Stephens et al. (2000), Evanno, Regnaut et al. (2005) and Rosenberg (2002).

The study area was divided into 13 geographic regions (Fig. 2, Table 3), roughly corresponding to the regions proposed in Hultgård (1987). Based on the original allocation of individuals to the four Structure clusters, the genetic structure of populations and regions were calculated in the following way: The allocation to each cluster for all the individuals in a population was summed up. The total allocation to one cluster for all the individuals was then divided by the total allocation to all four clusters, resulting in the population's fraction of allocation to this one cluster. When this fraction was higher than 60%, the population was assigned to this cluster. Populations that could not be assigned to any cluster according to this criterion were set as mixed (Table S1 in Appendix). The fraction of allocation to each cluster was also calculated for each region and the genetic structure of regions was visualised using pie charts. In the following, the phrase "genetic structure" will refer to this overall allocation of a population or a region.

Similarity between the AFLP multilocus phenotypes was calculated using Dice's similarity measure (Dice 1945) and visualised using principal coordinate (PCO) analyses in PAST ver. 2.08 (Hammer, Harper et al. 2001). Structure groups, geographical regions and ploidal levels were superimposed onto the PCO plots.

Intraregional diversities were calculated using Nei's gene diversity, D (Nei 1987):

$$D = \frac{n}{n-1} \cdot \left(1 - \left(freq(0)^2 + freq(1)^2 \right)\right)$$

for each marker, before calculating the region average diversity. Confidence intervals for diversity values were calculated by bootstrapping (Ehrich 2006). Rarity index, i.e. the sum of the occurrence of a marker in an individual divided by total number of occurrences of this marker, equal to "down-weighted markers" as described in Schönswetter and Tribsch (2005), was calculated for each region. Assuming that all markers were randomly distributed in all populations, a 95% confidence interval for the rarity index of each region was calculated. Values of rarity falling outside of these confidence intervals were considered significantly

large or small (Ehrich 2006). Regions with less than three individuals (Japan, the Netherlands) were included in the diversity analyses, but the results were interpreted with care.

Analyses of molecular variance, AMOVA (Excoffier, Smouse et al. 1992), between and within populations and groups were carried out in Arlequin ver. 3.5 (Laval and Schneider 2005) based on pairwise distances. Populations were grouped according to geographical regions, ploidy, and allocation to the Structure clusters. Mixed populations that did not allocate to one of the Structure clusters in more than 60% of the runs were excluded from all AMOVA analyses, leaving 35 populations (Table S1 in Appendix). Neighbour-joining networks were calculated using SplitsTree ver. 4.10 (Huson 1998; Huson and Bryant 2006) and uncorrected P as a measure of genetic distance (Huson 2005).

Growth experiment

According to previous studies and climatic data, diploid populations of *P. palustris* are distributed at lower latitudes with shorter days and higher mean temperatures during the growth season, while tetraploid populations are prevalent at higher latitudes with longer days and lower mean temperatures during growth season (E-klimatjenesten, Norwegian Meteorological Institute; Hultgård 1987), complicating the testing of a hypothesis concerning different ecological flexibility in di- and tetraploid *P. palustris*. However, if tetraploid *P. palustris* plants exhibit higher ecological flexibility than diploids towards varying temperature or day length, one would expect tetraploids to perform better than diploids also when treated with temperatures and day lengths that are deviating from that of their home environment. In order to examine whether tetraploid *P. palustris* displayed greater flexibility to varying day lengths and temperatures than diploid *P. palustris*, a null hypothesis was set up:

H_0: There is no difference in survival and growth rate between diploid and tetraploid *P. palustris* when treated with four different combinations of day length and temperature.

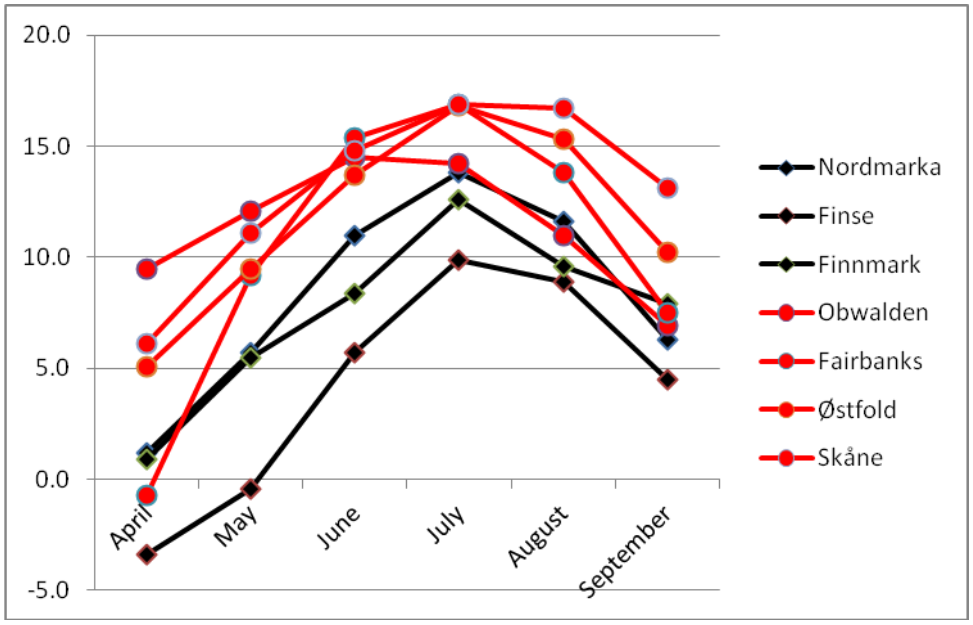

Figure 3. Mean temperatures from April to September, 2010 for the seven localities where seeds of *Parnassia palustris* for Experiment 1 and 2 were collected. Red line indicates diploid populations. Black line indicates tetraploid populations.

In order to exclude the effect of local adaptations, several populations from each ploidy level was included to test the null hypothesis. Four diploid populations were selected, representing a latitudinal range of 46.8 ºN – 64.8 ºN and mean temperatures in July between 14.2 – 16.9 ºC. Three tetraploid populations were selected, representing a latitudinal range of 59.9 ºN – 70.5 ºN and mean temperatures in July between 9.9 ºC and 13.8 ºC (E-klimatjenesten, Norwegian Meteorological Institute, Oslo; Table 2, Fig. 3).

Table 2. Populations and ploidal levels (estimated by flow cytometry) included in the growth experiment of *Parnassia palustris.* Latitudes (Lat.) are according to the coordinate system Euref 89. Mean temperatures (ºC) for each locality from April to September 2010 are collected (E-klimatjenesten, Norwegian Meteorological institute, Oslo).

						Mean temperatures April - September 2010					
Pop nr	**Country**	**Locality**	**Sublocality**	**Ploidy**	**Lat.**	**April**	**May**	**June**	**July**	**Aug.**	**Sept.**
4	Norway	Nordmarka	Karrussputten	4x	59.8994	1.2	5.7	11	13.8	11.6	6.3
8	Norway	Finse	Kvannjolnut	4x	60.6053	-3.4	-0.4	5.7	9.9	8.9	4.5
16	Norway	Finnmark	Durevuoppi	4x	70.4783	0.9	5.5	8.4	12.6	9.6	7.9
13	Switzerland	Obwalden	Engelberg	2x	46.8253	9.5	12.1	14.5	14.2	11	6.9
37	Alaska	Fairbanks	Sheep Creek	2x	64.8378	-0.7	9.2	15.4	16.9	13.8	7.5
38	Norway	Østfold	Butjern	2x	59.4874	5.1	9.5	13.7	16.8	15.3	10.2
66	Sweden	Skåne	Stängby mosse	2x, 3x*	55.7756	6.1	11.1	14.8	16.9	16.7	13.1

*Ind 66_5 and 66_6 are 3x.

Because of extremely poor germination and seedling survival in this experiment (Experiment 1), a second experiment (Experiment 2) was initiated with a higher number of replicates to strengthen the testing of the null hypothesis. Only populations 4 (Nordmarka, Norway) and 66 (Skåne, Sweden) were included in this experiment. These populations are located at similar latitudes, and experience similar day lengths, although the mean temperature in the growth season is diverging (Table 2).

To ensure that variation within populations was represented, seeds from five to ten plants from each population were pooled and germinated as previously described. After approximately three weeks of germination, surviving plants were placed in separate pots (8C OS plastic pots) with Herbia plant soil mixed with perlite in a 5:1 proportion. Environmental conditions were the same as during germination. Because of poor germination and survival, three series of germination were performed to provide enough plants for the experiments: Seeds were put to germinate 25. August, 20. September and 10. November 2010. Germination date was examined as a factor during the data analyses.

The growth chamber experiments took place in the Phytotron (UiO) from 19. January to 10. May 2011 (Experiment 1) and from 18. March to 27. June 2011 (Experiment 2). Before start, each leaf was compared to a graph paper to assess the surface in mm^2, and the diameter of each plant was measured from the root of the plant to the longest leaf in the rosette. Each plant was characterized as 1- Small (leaf surface under 4 mm^2 and less than 4 mm in diameter), 2 - Medium (leaf surface between 4 and 25 mm^2 and diameter between 4 mm and 10 mm) and 3 - Large (leaf surface more than 25 mm^2 and diameter more than 10 mm), and the size at the beginning of the experiment was examined as a factor during the data analyses.

For each treatment, six replicates, i.e. plants of similar size from the same population, were planned. However, because of poor survival, only five similarly sized replicates were possible for treatment of populations 4, 13, 38 and 66. For populations 37, 16 and 8, only 2, 1 and 3 similarly sized replicates, respectively, were given each treatment. The replicates from each population were spread out in a randomized pattern to avoid bias from the possibly varying microclimate in the growth chambers.

Experimental design

Factorial experimental design was used in Experiment 1 to study the effect of different day lengths and temperature regimes on growth and survival of di- and tetraploid populations. A combination of four different treatments was used: Short day conditions (SD; 12 h high light level and 12 h darkness), resembling day length conditions at lower latitudes; long day conditions (LD; 12 h high light level and 12 h dim light level), resembling day length conditions at higher latitudes; low temperature (LT; 15 ºC at daytime and 9 ºC at night time), resembling temperature at higher latitudes in the growth season, and high temperature (HT; 26 ºC at daytime and 20 ºC at night time), resembling temperature at lower latitudes in the growth season. In Experiment 2, only SD treatment was given in addition to LT and HT, as SD treatment probably was too stressful and resulted in low survival.

The climate factors were controlled by the climate computers (CWO Volmatic 296 computer, DGT Volmatic A/S, Vallensbæk Strand, Denmark). Only artificial light with daylight quality (Powerstar HQI-BT 400 W/D, OSRAM, Augsburg, Germany), giving a photosynthetic photon flux density (PPFD) of 200 ± 10 µmol m^{-2} s^{-1} at plant height, was used. Dim night time light for the long day treatment was provided by a fluorescent lamp (Luminette 58W/840, Aura, Karlskrona, Sweden) giving 0.6 – 3 µmol $m^{-2}s^{-1}$ at plant height.

Approximately every fourteenth day, variables describing growth were measured for each plant. This included diameter of the rosette and height of the highest leaf, both measured without stretching out stems or leaves. The number of leaves was counted, and total leaf surface was determined by visual registration of each leaf over a graph paper. At the first five times of measurements, a photo was taken of each plant including a graph paper, to be able to quality control the measurements; after this period (70 days), the plants were too large for a picture to provide any useful information. Number and approximate date of flower buds were recorded.

When the largest plants were approaching a size where their subsequent growth could be limited by pot size and lack of nutrition, the experiments were closed. Final leaf number and leaf surface were recorded for each plant. Each plant was harvested, placed in aluminum bags and dried for >48 hours at 105 ºC , using a Sartorious CP 423 S balance (Sartorius AB, Gottingen, Germany) with 0.001 g precision.

Generalised Linear Models

Many of the basic statistical methods assume that the response variables are continuous and normally distributed, and that the variances of the response variables are constant (Field 2005; Crawley 2007; Moran, Solomon et al. 2007). The present study included both discrete and continuous response variables, that either were having a Gamma-like distribution (i.e. leaf number, leaf surface and dry weight; Fig. 9), or were binominal data (i.e. survival). In both of these cases, the variation cannot be assumed to be constant (Crawley 2007). A Generalised Linear Model (GLM) can be used to model response variables with non-linear means, e.g. binary data, via a link function. The distribution of the response variables can be specified in the GLM, and it can thus be applied on response variables with non-normal distributions, and therefore without assuming that the variance is constant (Crawley 2007; Moran, Solomon et al. 2007; Dalgaard 2008).

To identify predictors and interactions between predictors giving an effect on each of the response variables, GLM was computed, allowing two-way interactions between predictors, using the glm functions in the statistical software R v. 2.12.2 (Crawley 2007; Dalgaard 2008). The null hypothesis H_0: Intercept = 0 was tested for each predictor and a standard p-value of 0.05 was applied. Due to very low survival rate and, consequently, lack of data, the Size:Population interaction and the Germination date:Size interaction had to be excluded from the full model. Preliminary analyses showed that these interactions were not explaining much of the variance in the response variables (not shown). Only plants from Experiment 1 were included in the GLM analyses.

Various error distributions can be specified in a GLM, depending on the nature of the error of the response variables (Crawley 2007). To determine which distribution to use, it is of interest to find k by solving the equation

1) $\mathrm{Var}(y) = a\,\mu^k$

To do this, the variance of the response variables was examined by comparing with a model assuming a Poisson distribution (Pers. Comm., Storvik 2011). The data were split into intervals, the variance within these intervals was estimated, and equation 1 was solved to find an approximation for k (not shown). For a gamma distribution, $\mathrm{Var}(y)^2$ is proportional to the mean when the shape parameter is held constant, i.e. k = 2 in equation 1, while for the inverse Gaussian distribution, $\mathrm{Var}(y)^3$ is proportional to the mean when the shape parameter is held

constant, i.e. k = 3 in equation 1 (Crawley 2007). According to the estimates of k, a gamma error distribution was therefore chosen to model leaf number (k = 1.997) and leaf surface (k = 1.86). For the GLM predicting dry weight an Inverse Gaussian error distribution was chosen (k = 2.906). A binomial error distribution is suitable for data on proportions (Crawley 2007), and was chosen for the GLM predicting the survival data.

A logarithmic link function is used to model responses when the dependent variable is assumed to have a logarithmic relation to the predictors (Crawley 2007). A logarithmic link function was employed to model the leaf number, leaf surface and dry weight. A logit link function (log (p/1-p)) transforms the data from probability data, belonging to the binomial probability distribution, to a linear scale (Wilson and Hardy 2002; Crawley 2007), and was applied for modelling plant survival.

As opposed to the Akaike Information Criterion (Akaike 1974), the Bayesian Information Criterion (BIC) takes the number of observations into account when calculating the penalty for adding a predictor (Schwarz 1978). BIC was employed to find the most parsimonious and informative model by excluding the least informative predictors from the full model, which included the predictors Day Length, Ploidy/Population, Germination date, Size and Temperature.

The following model-checking plots were employed to assess the previously described assumptions of variance and error structure (all in Crawley 2007). Residuals against fitted values, and scale-location (square root of the standardized residuals against the fitted values), both displaying if there is a trend of increasing or decreasing residuals with increasing fitted values. Q-Q-plots, showing how well the observed residuals were corresponding to the predicted residuals, and Cook's distance and residuals vs. leverage (standarized residuals as a function of leverage), both highlighting the y values that have the biggest impact on the predictor estimates.

Chi-squared tests were performed to find the p-values for the null hypotheses: H0: Model (Table S3 in Appendix). A large p-value (> 0.05) indicated that there was no reason to reject the null hypothesis, i.e. the model (Pers. Comm., Storvik 2011; Crawley 2007).

Survival analyses: Cox' regression

Survival analyses can handle censored cases; i.e., plants that either have missing data, plants that survived until the end of the experiment, and plants that started treatment at different times (Cox 1972; Dalgaard 2008). To be able to use both surviving and dead plants from Experiment 1 and Experiment 2 in the same analyses, parametric survival analyses using Cox' regression were employed. The risk of failing at a specific time, i.e. the hazard rate, is defined as the probability of a plant dying in the "next instant" or next time frame, given that this individual has survived until this time (Barlow, Marshall et al. 1963). The Cox' regression model of proportional hazard assumes that the underlying hazard rate is proportional, i.e. that it is a function of the predictors. Cox' regression can be used to model data sets with non-constant hazard and censored cases (Cox 1972; Crawley 2007). Kaplan-Meier curves can be used to estimate survival functions, i.e. the probability of being alive at a given time, from censored data (Dalgaard 2008) and was used to visualise the survival function and hazard rate as a result of time.

BIC (Schwarz 1978) was used as an information criterion to discard uninformative predictors from a maximal model. Experiment number was introduced as a predictor in the model testing, and Population/Ploidy were tested separately due to their high correlation. Kaplan-Meier curves were calculated for each predictor. The full model thus included the predictors Day Length, Ploidy/Population, Germination date, Size, Experiment number and Temperature. Two-way interactions were included, except the following interactions: Experiment number:Size, Experiment number:Temperature, Size:Population and Germination date:Size. These interactions were excluded due to missing data.

For the predictors, where the test of proportional hazard showed that the assumption of proportional hazard was violated, a non-parametric log-rank test (Dalgaard 2008) was applied to find the influence of the individual predictors without assuming proportional hazard.

Results

Ploidy

Ploidal levels were successfully estimated with flow cytometry for 385 plants from 58 populations (Table S1 in Appendix; Fig. 1). The results obtained from analyses of silica dried material were identical to the results obtained from fresh material in the 30 samples analyzed using both these approaches. The distribution of the cytotypes generally confirmed results from previous studies, where the diploids are found to be dominating in Southern Scandinavia and southwards to the Central and Southern European mountains, while the tetraploids are found mainly in Northern Fennoscandia (Hultgård 1987). However, an exception to this pattern was detected. Population 49 from Slovenia and population 50 from Austria turned out to be mainly tetraploid. The network analysis including only populations from the Alps and Slovakia (Fig. 4) showed that individuals from the tetraploid population 50 (Austria) clustered together with all but one individual from the tetraploid population 49 (Slovenia), constituting an "Eastern Alps" tetraploid clade. One Slovenian tetraploid individual clustered with diploids from Austria and Slovenia.

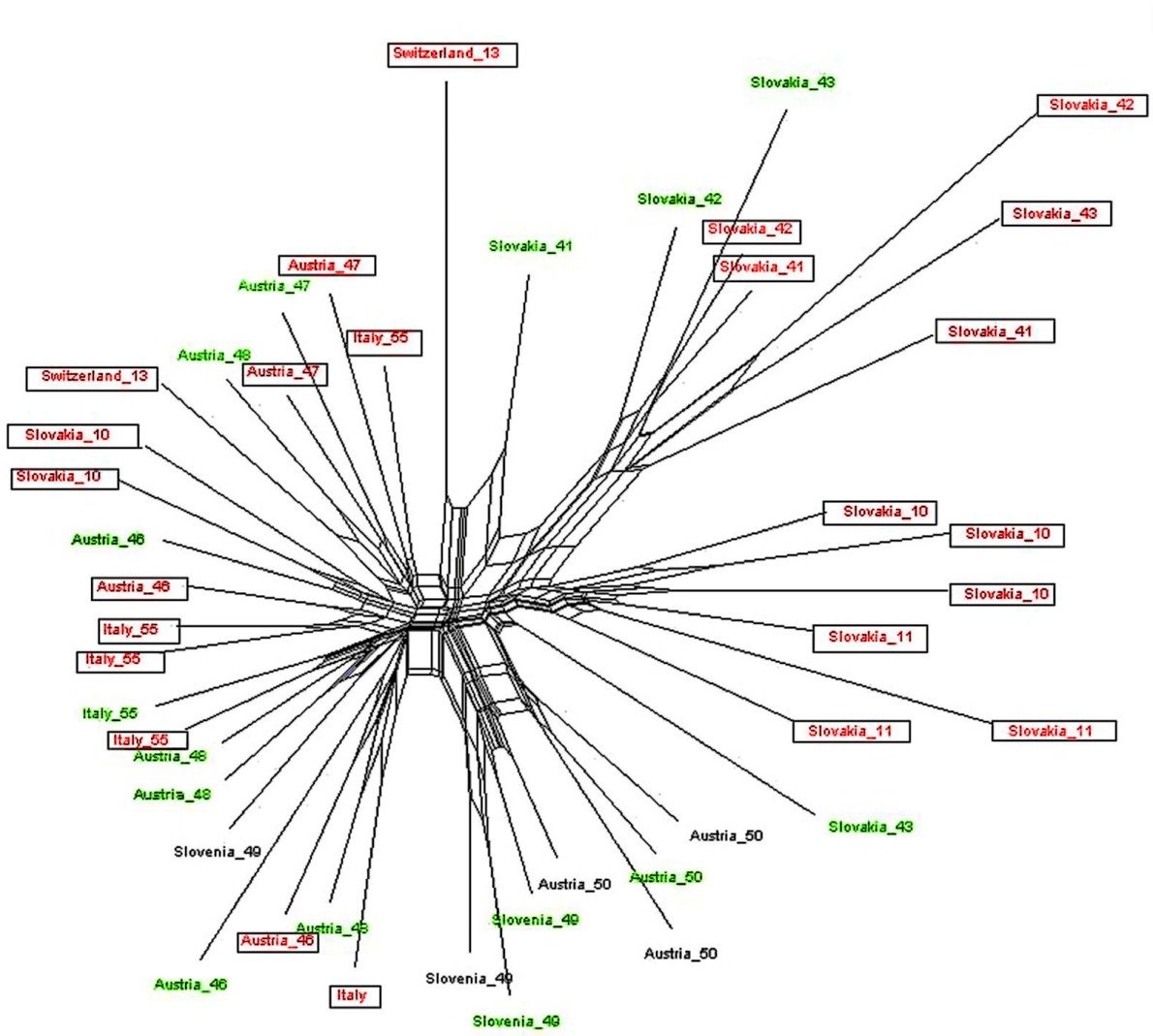

Figure 4. Neighbor-joining network of Central European individuals of *Parnassia palustris* based on 144 AFLP markers, using uncorrected P as a measure of genetic distance. The cytotype of the individual is indicated with the following colors: Red – diploid, black – tetraploid, green – unknown.

Di- and tetraploid populations were growing in sympatry, i.e. their distribution ranges were overlapping, in two areas: Southern Norway and Sweden and the Alps. In five populations, more than one ploidal level were detected. The mainly diploid population 10 from Slovakia comprised one tetraploid individual, whereas population 32 from the Netherlands was diploid except for one pentaploid individual. A triploid was found in the otherwise diploid Slovakian population 43, and this was also the case for population 66 from Southern Sweden. A pentaploid individual was found in the Austrian population 50.

AFLP analyses

A total of 162 individuals from 51 populations was successfully analysed for AFLPs, of which 158 were scored for 144 markers and four were scored for 126 markers (Table S1 in Appendix). The technical difference rate, calculated as the number of differences per duplicated profile divided by the total number of fragments per profile (Bonin, Bellemain et al. 2004), was 4.32%. Fragment number per individual ranged between 20 and 60, and average fragment number per individual was 41 and 45 for diploids and tetraploids, respectively. Private fragments were found in Japan, Alaska and alpine Central Norway, holding five, three and one private fragment, respectively.

BAPS analyses gave highest log marginal likelihood (-8417.0) when the AFLP data were partitioned into six clusters (not shown), two of which consisted solely of individuals from Japan and Alaska. In order to obtain better resolution of the genetic structure in the area of main interest, only the European samples were included in the Structure analyses, where Ln P(D), similarity coefficients (Fig. S1 in Appendix) and ΔK (not shown) indicated a partitioning into four clusters. These clusters, hereafter named cluster 1–4, corresponded well to the clusters identified by BAPS; only five of the individuals, which allocated to a Structure cluster in more than 60% of the runs, were placed in a different cluster by BAPS. However, most individuals showed mixed allocation to the four Structure clusters (Fig. S2 in Appendix). When overall allocation to each cluster was calculated, most of the regions displayed mixed genetic structure, although regional differences in the genetic structure could be identified (Fig. 2).

A high degree of allocation to Cluster 1 characterized individuals from Northern Fennoscandia and the East region, comprising Balticum and Western Russia (Fig. 2). To a certain extent, individuals from both lowland and alpine areas of Central Scandinavia also allocated to this cluster. Cluster 2 was clearly dominating in the Alps, and to a lesser degree also in Slovakia, Southern Sweden and Norway, Iceland, the lowlands of Central Scandinavia and Denmark. The cluster was poorly represented in Scotland, Northern Fennoscandia and the East region. Cluster 3 had its main centre in North-Western Europe, dominating in populations from Denmark and Scotland. Populations from Southern Norway and Sweden also displayed a high degree of allocation to Cluster 3, as did the alpine areas of Central Scandinavia. Despite its geographical position, the Netherlands showed only little connection

to this North-Western Cluster 3, whereas in Slovakia, which is situated far from North-Western Europe, Cluster 3 constituted almost half of the region's genetic structure. Northern Fennoscandia, the East region and the Alps had little or no allocation to this cluster. Cluster 4 clearly dominated in the Netherlands and was also characterizing the genetic structure of Northern Fennoscandia and, to a lesser extent, the Icelandic population. The other regions showed little or no allocation to this cluster.

Table 3. Geographical regions used in analysis of population structure of *Parnassia palustris*, and diversity measures for each region, based on 144 AFLP markers. Confidence intervals (CI) for Nei's gene diversity (D) values are calculated by bootstrapping over the markers. Using permutation tests, a 95% confidence interval is calculated for the rarity (R) of markers of each region, assuming that the markers were distributed randomly. When the rarity value for a region lies outside of this 95% confidence interval, the value is considered significantly large or small.

Abbreviation	Region	Populations	Sample size	Rarity R	CI R low*	CI R high**	Diversity D	CI D low*	CI D high**
ISL	Iceland	1, 2	5	0.86	0.65	1.13	0.19	0.15	0.24
NFSc	Northern Fennoscandia, north of 66 °N	15-28	42	0.87	0.79	0.93	0.24	0.21	0.27
CSc	Lowlands of Central Scandinavia between 66 °N and 61 °N	52-54	10	0.68	0.71	1.04	0.18	0.15	0.21
CN alp	Central Norway >700 m asl	30, 12, 8, 9	7	0.89	0.68	1.1	0.21	0.17	0.25
SNS	Sweden and Norway south of 61°N	3-5, 7, 38, 61, 66	14	0.85	0.72	1	0.2	0.16	0.23
DNK	Denmark	57, 59	10	0.6	0.71	1.05	0.17	0.14	0.21
GBR	Scotland	62, 68	6	0.81	0.65	1.11	0.15	0.12	0.19
NLD	The Netherlands	32	2	0.67	0.56	1.34	0.22	0.15	0.29
Alps	Switzerland, Austria, Italy, Slovenia	13-14, 46-50, 55	27	0.62	0.76	0.96	0.15	0.12	0.18
SVK	Slovakia	41-43, 10-11	17	0.64	0.73	1	0.2	0.17	0.23
East	Balticum/Western Russia	44, 51, 60	11	0.85	0.71	1.03	0.19	0.15	0.22
AL	Alaska	37	5	2.5	0.65	1.12	0.23	0.19	0.27
JA	Japan	67	2	3.65	0.56	1.3	0.11	0.06	0.17

* Lower bound, 95% Confidence interval

** Upper bound, 95% Confidence interval

In Southern Norway and Sweden, the diploid populations 38 and 66 and tetraploid populations 7 and 6 were relatively closely located. In the Alps, the diploid populations 13, 14, 46, 47, 48, 49 and 55 and the tetraploid populations 49 and 50 were also closely located. These populations are hereafter referred to as sympatric. The genetic structure of the diploid populations from Southern Sweden diverged only slightly from that of the two tetraploid populations. The genetic structure of the tetraploid populations from the Alps was very similar to that of the sympatric diploid populations (Fig. 5).

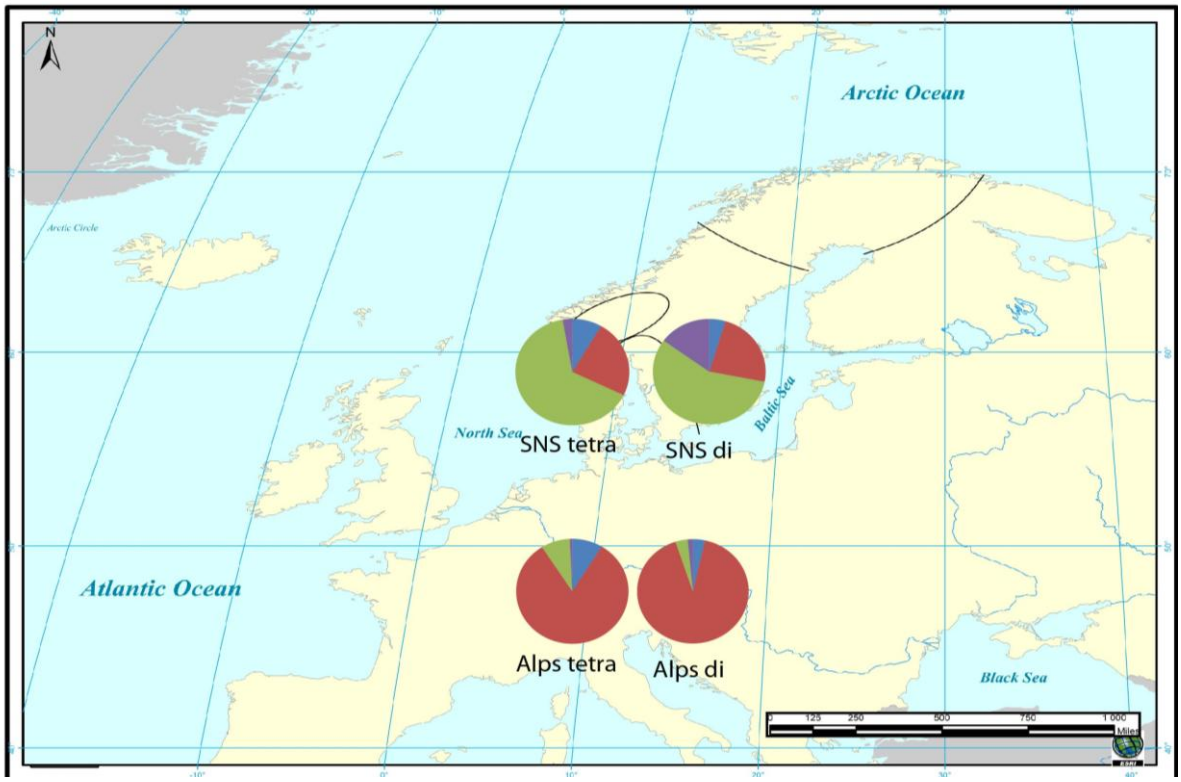

Figure 5. Allocation analyses based on 144 AFLP markers grouped the European individuals of *Parnassia palustris* into four clusters: Cluster 1 (blue), Cluster 2 (red), Cluster 3 (green), and Cluster 4 (violet). The pie charts show the genetic structure (i.e. the overall allocation to each of the four clusters for all individuals) in di- and tetraploid populations from the sympatric areas in Southern Norway and Sweden (SNS; including the diploid populations 38 and 66 and the tetraploid populations 7 and 61) and the Alps (including the diploid populations 13, 14, 46, 47, 48 and 55 and the tetraploid populations 49 and 50).

The complex genetic structure characterizing most of the regions was reflected in the ordination analyses. PCO of the total dataset displayed distinct Japanese and Alaskan clusters, both clearly separated from the rest of the individuals (not shown). When excluding these populations from the analyses, considerable overlap between the regions was displayed (Fig. 6).

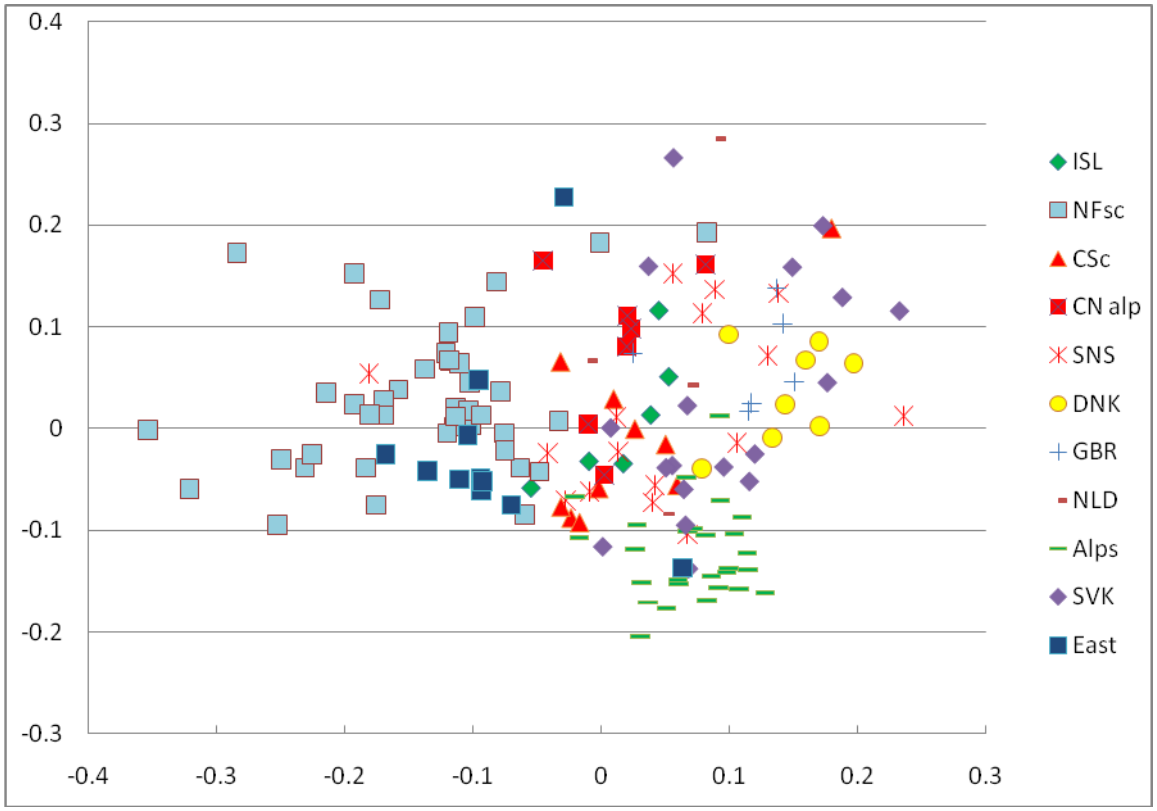

Figure 6. PCO analyses of European samples of *Parnassia palustris,* based on 144 AFLP markers using Dice as a distance measure. Individuals are assigned to the following regions, defined in Table 3: Iceland (ISL), Northern Fennoscandia (NFSc), Central Scandinavia (CSc), Central Norway alpine areas (CN alp), Southern Norway and Sweden (SNS), Denmark (DNK), Great Britain (GBR), the Netherlands (NLD), the Alps (Alps), Slovakia (SVK), and Western Russia and Balticum (East).

However, component axis 1, explaining 14.3% of the variation, represented a gradient from north to south. Individuals from Northern Fennoscandia and the East region were found at the left end, while individuals from Central Scandinavian lowlands and alpine areas, and also Icelandic individuals, occupied an intermediate position. Individuals from further south on the European continent were found at the right end of component axis 1. Component axis 2, explaining 10.3% of the variation, partly separated individuals from the Alps from the rest. In order to examine how the clusters defined by the Structure analyses relate to each other, individuals were allocated to a Structure defined cluster if the analyses allocated it to this cluster in more than 60% of the runs (Fig. S3 in Appendix). Cluster 1 (mainly characterising Northern Fennoscandia and the East region), Cluster 4 (mainly characterising the Netherlands and Northern Fennoscandia) and Clusters 2/3 (mainly characterising the Alps and North-Western Europe, respectively) were more or less separated along component axis 1. Component axis 2 further separated Cluster 2 (mainly characterising the Alps) and Cluster 3

(mainly characterising North-Western Europe). Although Component axis 1 tended to separate diploids and tetraploids, there was a pronounced overlap between cytotypes (Fig. 7).

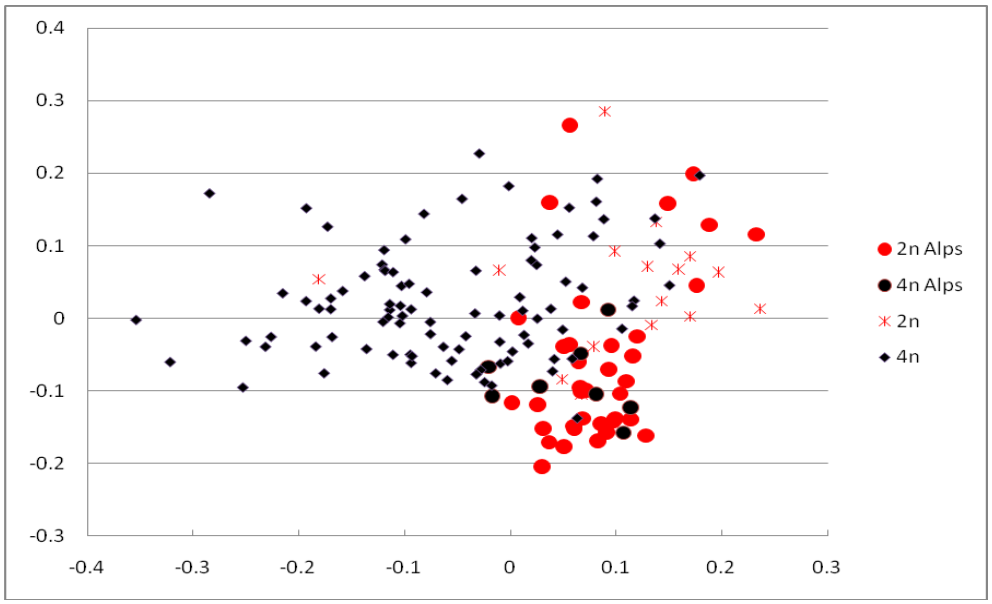

Figure 7. PCO analyses of European samples of *Parnassia palustris*, based on 144 AFLP markers using Dice as a distance measure. Diploids from the Alps, i.e Austria, Italy, Slovenia and Switzerland (2n Alps), diploids from the other localities sampled (2n), tetraploids from the Alps (4n Alps), and tetraploids from the other localities sampled (4n) are marked on the plot.

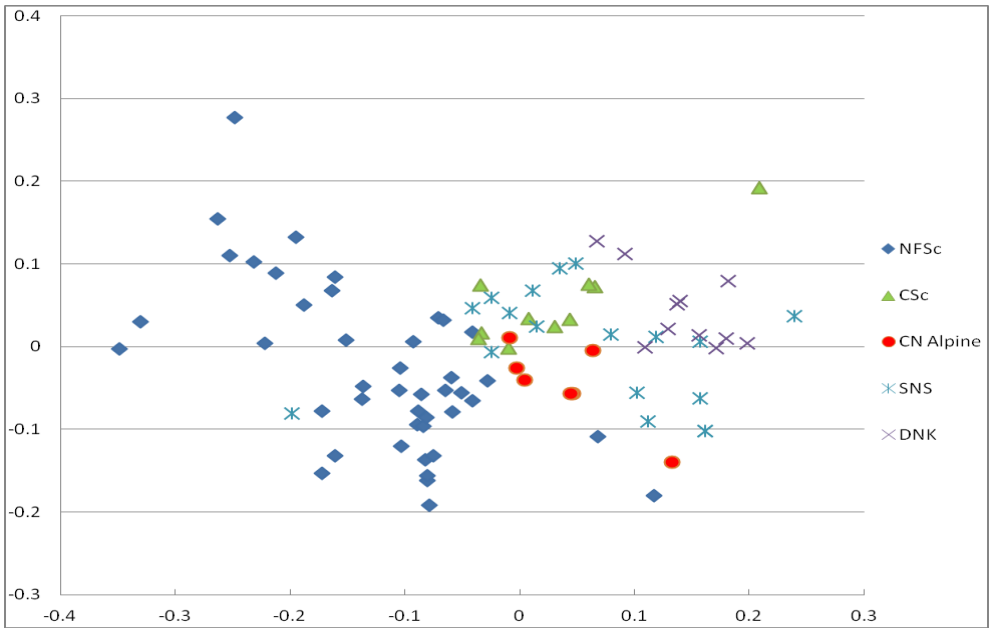

Figure 8. PCO analyses of Fennoscandian samples of *Parnassia palustris*, based on 144 AFLP markers using Dice as a distance measure. Individuals are assigned to the following geographical regions: Northern Fennoscandia (NFSc), Central Scandinavia (CSc), Central Norway > 700 m a.s.l. (CN alp), Southern Norway and Sweden (SNS), and Denmark (DNK) as defined in Table 3.

The tetraploids from the Alps seemed to group more closely with the diploids from the Alps than with tetraploids from further north; this was also the case for the tetraploids from the sympatric area of Southern Sweden (Fig. 7). When including only individuals from Fennoscandia in the PCO analysis (Fig. 8), component axis 1 explained 13.7% of the variation and more or less separated individuals from Central (Central Norway between 66 °N and 61°N), Alpine (Central Norway > 700 m a.s.l.), South (Southern Norway and Sweden), and Denmark, from individuals from North (Fennoscandia north of 66 °N). Component axis 2 explained 5.2% of the variation, and mainly accounted for variation within each region. Component axis 3 did not contribute additional information in any of the PCO analyses. Compared to other regions, Alaska and Japan had rarity values higher than what was expected by chance, given a confidence interval of 95% and assuming a random distribution of markers (Table 3). Slovakia, the Alps, Central Scandinavia and Denmark had rarity values lower than what was expected by chance (Table 3). The average genetic diversity over all regions was 0.16; this value was not significantly altered when only European individuals were included. Genetic diversity was highest in Northern Fennoscandia (0.24), Alaska (0.23) and the Netherlands (0.22), while it was lowest in Japan (0.11), the Alps (0.15) and Scotland (0.15). The pattern of genetic diversity was thus to a certain degree corresponding with the genetic structure revealed in allocation analyses. The East and Northern Fennoscandian regions, which had high degree of allocation to Cluster 1, had higher diversity than the average. The Netherlands, characterized by Cluster 4, also exhibited high diversity. Denmark and Scotland, characterized by a high Cluster 3 allocation, seemed to comprise lower diversity than the average, and so did the Alps, with a high degree of Cluster 2 allocation. Cluster 1 and 4, thus, seemed to comprise regions with higher genetic diversity than average, while Cluster 2 and 3 comprised regions with less genetic diversity.

Table 4. Nei's gene diversity, D, based on 144 AFLP markers, in diploid and tetraploid individuals of *Parnassia palustris*. Confidence intervals (CI) for the diversity values are calculated by bootstrapping over the markers.

Ploidy	Sample size	Diversity D	CI D low*	CI D high**
2x	58	0.22	0.20	0.25
4x	100	0.23	0.21	0.26

* Lower bound, 95% Confidence interval

** Upper bound, 95% Confidence Interval

Only a slight difference in average genetic diversity could be detected between cytotypes; diploids and tetraploids held an average diversity of 0.22 and 0.23, respectively (Table 4). No significant difference in rarity of markers could be detected between cytotypes. AMOVA analyses (Table 5) of the 35 populations that could be placed in a Structure cluster assigned 67.93% of the genetic variation to the within-population component. Variation between the two main ploidal levels accounted for only 4.88% of the total genetic variation. When individuals were grouped according to the geographical regions shown in Fig. 2, variation among groups was 14.44%, reducing the within-population component to 65.75%. The highest level of among-group variation was obtained when the populations were grouped according to the Structure clusters. This partitioning reduced the within-population component to 64.93%, and accounted for 17.86% of the variation among groups. All values of variation within and among groups and populations were significant ($p < 0.05$).

Table 5. Analyses of molecular variance (AMOVA) in *Parnassia palustris*, based on 144 AFLP markers. AMOVAs are performed among populations, within populations and among groups. Populations are assigned to groups according to dominating ploidal level (2x or 4x), geographical regions as defined in Table 3 or Structure-defined clusters (Table S1 in Appendix); 14 populations that could not be assigned to one of the Structure-defined clusters (the population's fraction of allocation to any cluster was less than 60%) are excluded from all AMOVA analyses. Degrees of freedom (d.f.) are included.

	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
a) Populations	Among populations	38	1026.104	5.295	32.07
	Within populations	78	874.717	11.214	67.93
b) Ploidal level	Among groups	1	77.558	0.826	4.88
	Among populations within groups	37	948.546	4.869	28.80
	Within populations	78	874.717	11.214	66.32
c) Geographical regions	Among groups	9	439.538	2.439	14.44
	Among populations within groups	30	626.9	3.345	19.81
	Within populations	82	910.317	11.101	65.75
d) Structure clusters	Among groups	3	328.287	3.084	17.86
	Among populations within groups	35	697.817	2.973	17.21
	Within populations	78	874.717	11.214	64.93

[a] P-value < 0.05 for all estimations

Neighbor-joining networks including all individuals (Fig. S4 in Appendix) as well as only Fennoscandian individuals (not shown) confirmed the lack of major splits between ploidal levels, regions or populations. Although diploid individuals from the Alps tended to cluster together, as did tetraploid individuals from Fennoscandia north of 61 °N, no clear splits between the cytotypes could be detected. Diploids from Southern Norway and Sweden and Slovakia clustered together with tetraploid individuals from the Northern Fennoscandia. Tetraploid individuals from the East region clustered with diploid individuals from Southern Norway and Sweden. Individuals from the tetraploid population 50 (Austria) clustered with one individual from a diploid population from Slovakia. Out of four individuals from the tetraploid population 49 (Slovenia), two clustered with tetraploid individuals from Northern Fennoscandia, one clustered with mainly tetraploid individuals from population 50 (Austria) and one clustered with diploid individuals from the Alps, Denmark and Slovakia (Fig. 4).

As the Netherlands region consists of only two individuals, the results for this region will not be further discussed.

Growth experiment

Experiment 1 was ended after 109 days. Out of 184 plants included, 42 survived until harvest, giving a survival rate of 22%. Experiment 2 was ended after 99 days. Out of 38 plants, 31 survived until harvest, giving a survival rate of 69%. Only two plants survived the short day treatment; these were from the diploid populations 66 (Southern Sweden) and 13 (Switzerland). These two plants, and plants that had zero surface or zero leaves at the time of harvest, were excluded from the GLM analyses, but included in the survival analyses.

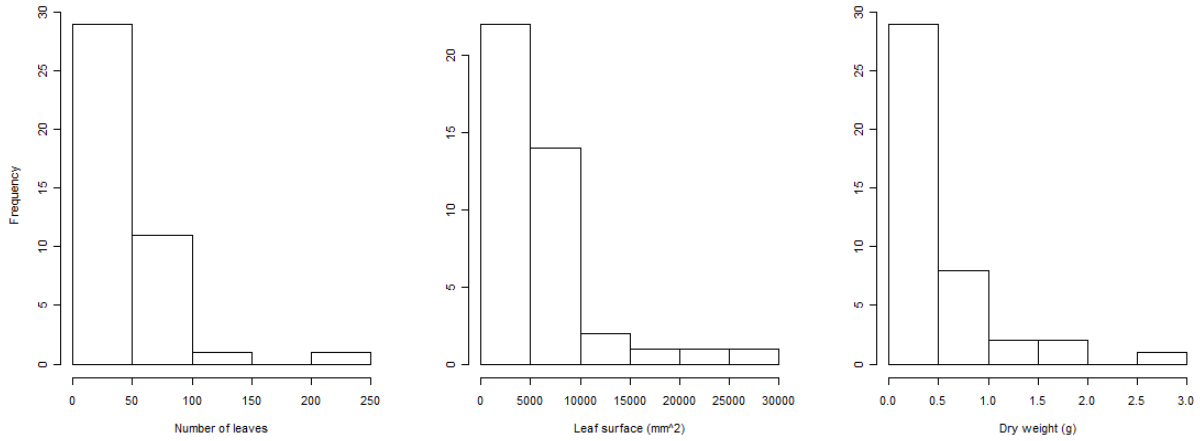

Figure 9. Frequency distributions of number of leaves, leaf surface (mm2) and dry weight (g) measured in the 42 plants of *Parnassia palustris* that survived until the end of Experiment 1. The plants were exposed to different temperature and daylight treatments during the 109 days long growth experiment.

Generalised Linear Models

The distribution of growth variables were not normally distributed (Fig. 9). The predictors Ploidy and Population were highly correlated, and were therefore analysed separately. Germination date and Size were also correlated, and whenever a model ended up including one of these predictors, it was replaced by the other to see how this affected the information criterion. The p-values, i. e. the probabilities for observing a value this large or larger if the null hypothesis (H_0: Intercept = 0) was true (Crawley 2007), were used to find predictors significantly influencing the response variables. Germination date was the most robust predictor for the number of leaves (p-value = 0.7627; Fig. 10, Table S3 in Appendix). Adding the predictors Temperature and Size to the model gave only a slight increase in BIC. The Residuals vs. Predicted values, the Q-Q-plot and the Scale-Location plot indicated that the Gamma error distribution and logarithmic link were appropriate for the number of leaves (Fig. S6 in Appendix). Size and Temperature were the most robust predictors for the leaf surface (p-value = 0.6083; Fig. 11, Table S3 in Appendix).

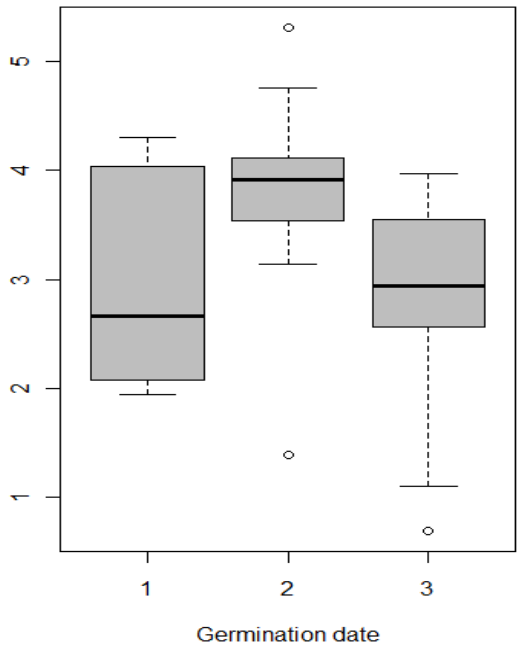

Figure 10. Germination date was the most robust predictor for leaf number (y-axis) in *Parnassia palustris* in the 109 days long Experiment 1. Plants from the following three germination dates were included in Experiment 1: 1. – 25. August (1), 2. – 20. September (2) and 3. – 10. November (3) 2010.

The high temperature treatment gave a significant increase in leaf surface (coefficient 0.7831, p - value = 0.027). The Residuals vs. Predicted values, the Q-Q-plot and the Scale-Location plot indicated that the Gamma error distribution and logarithmic link and were suitable for the leaf surface (Fig. S7 in Appendix).

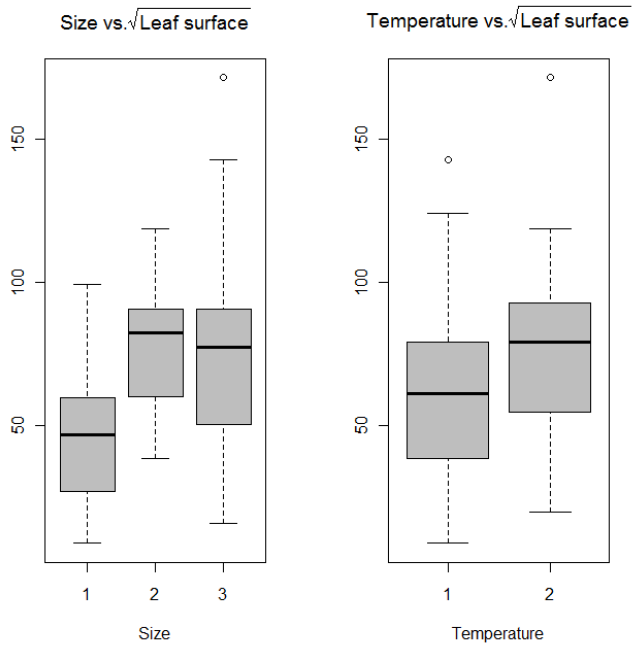

Figure 11. The most robust predictors for the leaf surface in *Parnassia palustris* were the size of the plants at the experiment start (Size) and the temperature treatment (Temperature). At the start of the experiment, the size of the plants was categorized as 1 - small, 2 - medium or 3 - large. The plants were divided in two groups that were treated with 1 - low and 2 – high temperature, respectively. Leaf surface was measured in mm2.

To predict dry weight, the most informative and parsimonious model included the predictors Germination date, Size and Temperature (p-value = 0.690; Fig. 12, Table S3 in Appendix). The high temperature treatment had a significant, positive effect on dry weight (p = 0.03797). The Residuals vs. Predicted values, the Q-Q-plot and the Scale-Location plot indicated that the inverse Gaussian error distribution and logarithmic link and were suitable for the dry weight (Fig. S5 in Appendix). The most informative GLM for survival data included only the Size predictor (p-value = 0.75; Table S3 in Appendix). The Q-Q-plot showed that the binominal error distribution was poorly fitted to the observed error distribution, as the line is far away from being straight (Fig. S8 in Appendix).

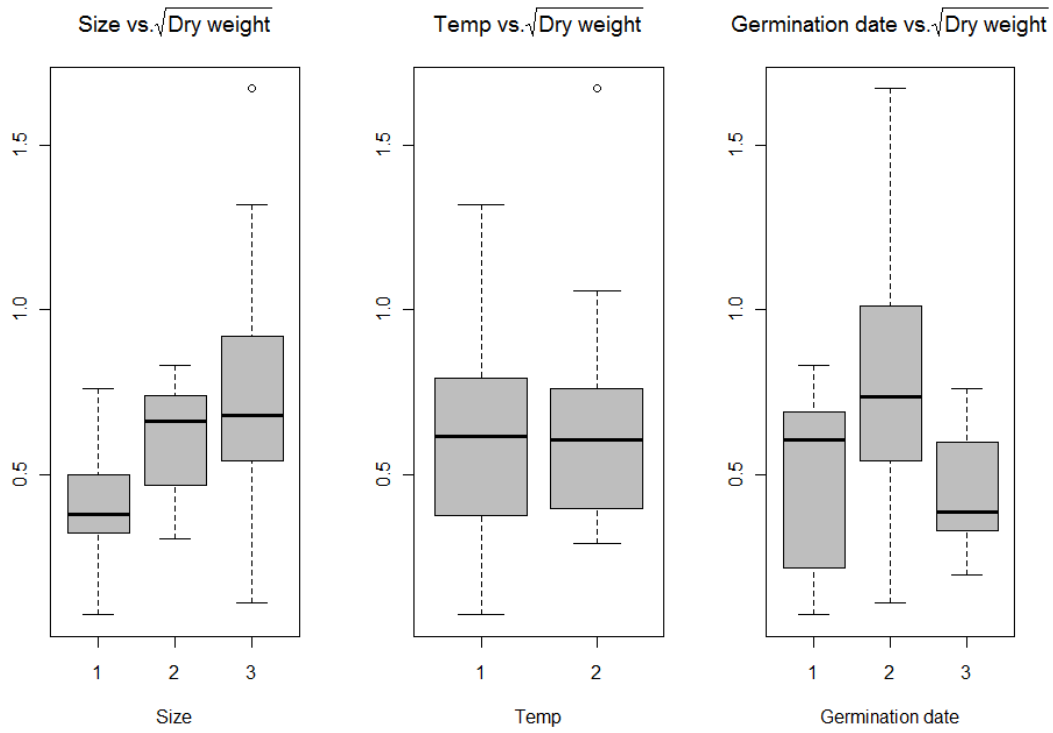

Figure 12. The most robust predictors for the dry weight of plants of *Parnassia palustris* that survived to the end of Experiment 1 were Germination date, the size of the plants at the experiment start (Size), and the temperature treatment (Temp; close to significant). Plants from the following three germination dates were included in Experiment 1: 25. August (1) , 20. September (2) and 10. November (3) 2010. The plants were divided in two groups that were treated with 1 - low and 2 – high temperature, respectively. At the start of the experiment, the size of the plants was categorized as 1 - small, 2 - medium or 3 - large.

Survival analyses: Cox'regression

A Cox' regression parametric model, assuming proportional hazard between the predictors, was employed to find the predictors that influenced the hazard rates. When including Ploidy as a factor, the most informative model according to the BIC included the predictors Experiment number, Day Length, Ploidy, Size, Temperature and an interaction between Ploidy and Day Length (R squared = 0.414; Table S4 in Appendix). Overall, tetraploid plants showed a non-significant tendency towards having lower hazard rate than diploids (coefficient = -0.276; p-value = 0.28) while tetraploids had significantly higher hazard rate than diploids during the high temperature regime (coefficient = 0.852; p-value = 0.013). When analysed separately, neither Ploidy nor Temperature showed a significant effect on survival. When including Population as a factor, only Experiment number, Day Length and Size were significantly affecting survival (R squared = 0.395; Table S4 in Appendix).

The assumption of proportional hazard was violated for the predictors long day and size 3 (Table S5 in Appendix). According to the Kaplan-Meier curves, the plants that were exposed to the short day treatment experienced a defined drop in survival between day 20 and 40, while plants that were exposed to long day treatment showed no such drop. Small plants showed a drop in survival between day 0 and 20, while medium-sized and large plants had no such drop. The effect of size on hazard rate decreased with time; the opposite was the case for the long day treatment (Figs. S9 and S10, respectively, in Appendix). Experiment number was also close to breaking the assumption of proportional hazard (Fig. S11 in Appendix, Table S5 in Appendix). When analysing these factors separately without assuming proportional hazard, using a non-parametric log rank test on the influence of each factor, Experiment number, Size and Day Length turned out to be the only factors significantly influencing the survival (not shown). However, the statistical power of the non-parametric log rank method is poor compared to the parametric Cox' regression. The effect of the violation of the proportional hazard assumption is not considered to be of great importance for the results, as the analyses work on an average effect scale (Pers. Comm., Reitan 2011). The model resulting from the Cox' regression will therefore be further discussed, with care however.

Discussion

Based on the results previously described, I will in the following address the five main questions that were raised in the introduction. I will discuss whether the genetic analyses of *P. palustris* support species rank for the European tetraploid cytotype. I will then describe and discuss how the distribution of cytotypes revealed in the present study relates to that of previous studies. Further, I will discuss how hybridisation and recurrent origins could have contributed to the observed genetic diversity. I will discuss whether immigration from different source regions could have caused a contact zone in Fennoscandia, and how such a contact zone could have affected genetic structure and levels of genetic diversity of *P. palustris* populations in this region. Finally, I will evaluate whether tetraploid populations of *P. palustris* display higher survival and growth rate than diploids under a range of temperatures and day lengths.

European tetraploid *P. palustris* does not deserve species rank

Among the sampled European individuals of *P. palustris*, ploidal level explained a very little proportion (4.9%) of the genetic variation. In contrast, differentiation between cytotypes of *Juncus biglumis* L. explained 55.3% of the AFLP variation in Schönswetter, Suda, et al. (2007), which lead the authors to suggest rank as a cryptic species for the polyploid cytotype. The network and ordination analyses did to a very low degree separate the cytotypes within *P. palustris*. Sympatric populations with different cytotypes were genetically very similar. This lack of correspondence between genetic variation and ploidy could be caused by both hybridisation between the ploidal levels, and recurrent origins of the tetraploid cytotype. These processes will be discussed further in the next sections. The molecular data therefore suggested gene flow between cytotypes, which was further supported by the existence of tri- and pentaploids (see below). There is increasing evidence in several other taxa for gene flow between cytotypes rather than absolute reproductive isolation between ploidal levels (e.g. Lyrene, Vorsa et al. 2003; Brochmann, Brysting et al. 2004; Slotte, Huang et al. 2008).

Parnassia palustris has been subject to different taxonomical treatments, based on varying morphological, distributional and cytological criteria (summarized in Table 22 on page 106 - 107 in Hultgård 1987). Already in his Flora Lapponica, Wahlenberg (1812) described the variety *tenuis* as smaller and more slender than the variety *palustris*. In his description of

Scandinavian material, Erlandsson (1942; in Hultgård 1987) suggested species rank for a large-grown diploid variety, a large-grown tetraploid variety and the smaller tetraploid variety that previously had been described as variety *tenuis*. However, no formal taxonomical suggestion to change the taxonomic rank was proposed. Löve (1950) stated that the two cytotypes had different distribution and different morphology. In accordance with Erlandsson (1942), he distinguished between the diploid *P. palustris* L. em. Löve and the tetraploid *P. obtusiflora* Rupr. em. Löve. However, he did not treat *tenuis* as a species, but included it as a subspecies of *P. obtusiflora.* Based on morphological differences and chromosome number, Hultén (1971) suggested to split *P. palustris* into two subspecies that were thought to hybridize in areas with overlapping distribution. These are an arctic-alpine tetraploid subspecies *neogaea* Fernald. and an Eurasiatic diploid subspecies *palustris.* Based on morphological variation in a wide range of habitats in Scandinavia, Hultgård (1987) found that the two cytotypes were indistinguishable.

In Northern Europe, three morphological/cytological groups are currently recognized according to the Pan Arctic Flora (PAF; Elven 2007 onwards). These are tall, diploid plants, with a southern distribution; large tetraploid plants morphologically seemingly inseparable from the diploids, but with a more northern distribution; and small tetraploid plants with an alpine and northern European/Russian distribution. Based on morphology, Elven (2007 onwards) suggested that the latter group could be justified as a separate taxon.

Traditionally, autopolyploid cytotypes have been treated not as independent units, but merely as cytotype variation within their parental species (Thompson and Lumaret 1992). However, employing different species concepts might lead to diverging taxonomical conclusions. The biological species concept (Mayr 1942) could be used to justify species rank to the autotetraploid if the cytotypes are reproductively isolated from each other. Soltis, Soltis et al. (2007) argued that even though autopolyploids might not be morphologically distinct from their diploid progenitors, they might function as separate lineages and be reproductively isolated. They suggested using multiple species concepts, e.g. the evolutionary, taxonomic and biological, in the taxonomical treatment of autopolyploids, and that taxonomic rank should be considered independently for each autopoplyploid (Soltis, Soltis et al. 2007). Mallet (2007) stated that autopolyploids in plants should be treated as species distinct from their progenitors provided that they "*form clusters that are more abundant than intermediates*

formed by hybridization between them". European autotetraploid *P. palustris* is separated neither reproductively nor morphologically from its diploid progenitor, and the cytotypes are not forming separate clusters or lineages. It does not fulfil any of the criteria set by Mayr (1942), Soltis, Soltis et al. (2007) or Mallet (2007), and species rank for autotetraploid *P. palustris* cannot be justified.

If the small-grown tetraploid alpine/northern plants should constitute a separate taxon, as suggested by Elven (2007 onwards), we would expect them to be genetically more similar to each other than to the surrounding, large-grown tetraploid plants, according to the evolutionary and the biological species concept (Mayr 1942; Soltis, Soltis et al. 2007). In the network and ordination analyses, Central Norwegian alpine tetraploids clustered with tetraploids from Central Scandinavian lowlands and Southern Norway and Sweden, and there was no clear separation between tetraploids from Northern Fennoscandia and plants from Southern Norway and Sweden. Based on AFLP data, the small-grown, alpine/northern tetraploid morphotype thus does not represent a separate lineage, and there seems to be no reason to rank it as a separate taxon.

The lack of genetic differences that these conclusions rely upon could also be because the genetic markers used are not suitable for revealing the genetic structure present. However, other studies have used a similar number of AFLP markers to provide valuable information on genetic structure and recent history of various plant species (Ehrich, Gaudeul et al. 2007; Eidesen, Alsos et al. 2007; Schönswetter, Suda et al. 2007; Westergaard, Jørgensen et al. 2010). The fact that my AFLP data provide overall geographical structure that make sense, provides credibility to the results and conclusions drawn above and in the following discussion.

Distribution of ploidal levels: Tetraploid populations in the Alps

The distribution of *P. palustris* cytotypes obtained generally confirmed the distribution described in previous studies (summarized in Hultgård 1987): Diploids are dominating in Central and Southern Europe and northwards to Southern Scandinavia, while tetraploids are found mainly in Northern Fennoscandia. The distributions of the two cytotypes are overlapping in a zone in Southern Scandinavia. The flow cytometry data revealed, however,

two tetraploid populations in the Julian Alps, Austria (pop. 49), and in Karawanken, Slovenia (pop. 50). The populations were sampled in the alpine zone (1400 – 2100 m asl) and displayed low rarity in AFLP markers, as would be expected in a population that has immigrated post-glacially. The only tetraploid populations that previously have been recorded from Southern Europe are from the Mount Olympus region in Greece (Hultgård 1987), and only diploid individuals have previously been reported from the Alps (Pers. Comm. Tribsch 2011; Hultgård 1987; Dobes and Hahn 1997).

More than one ploidal level were detected in five of the populations (Fig. 1, Table 1). Pentaploid individuals were found in otherwise diploid (pop. 32; the Netherlands) and tetraploid (pop. 50; Austria) populations. Two triploid individuals were found in an otherwise diploid population (pop. 66; Southern Sweden), which was located in an area where di- and tetraploids overlap in distribution range (Fig. 1; Fig. 2 in Hultgård 1987). Previously, both di-, tri- and tetraploid individuals have been found in one single population from this area (Dalsland, UH 64; Hultgård 1987), and one pentaploid individual has been recorded from an otherwise tetraploid population in the same area (Dalsland, UH 69; Hultgård 1987). The findings of tri- and pentaploids are consistent with results from other studies of *P. palustris*, where ploidal levels of 2*x* - 6*x* have been reported (e.g Gadella and Kliphuis 1968; Hultgård 1987; Gornall and Wenthworth 1992; Funamoto, Kondo et al. 2006), and confirm that the study species include a variety of ploidal levels.

High diversity a result of multiple polyploid origins and/or hybridisation between cytotypes

Sympatric di- and tetraploid populations from Southern Norway and Sweden clustered together in the ordination and network analyses of the AFLP data, as did sympatric di- and tetraploid populations from the Alps. Allocation analyses further confirmed that sympatric di- and triploid populations of both regions have very similar genetic structure. This, together with lack of genetic similarity of tetraploids from the Alps and Southern Norway and Sweden, could be explained by recurrent, independent origins of the tetraploid cytotype in these two areas. One of the tetraploid individuals from Slovenia clustered with diploids from the Alps and not with the other tetraploid Slovenian individuals in the networks, indicating that the tetraploids from Slovenia might result from more than one polyploidisation event.

Hultgård (1987) and Borgen and Hultgård (2003) proposed at least two different origins of the tetraploid *P. palustris* cytotype in Fennoscandia, in late-glacial and early post-glacial time. In this period, environmental conditions were drastically changing in the area where the tetraploid cytotypes were proposed to origin (i.e. in South-Central Scandinavia and Eastern Fennoscandia; Hultgård 1987). It has been proposed that the frequency of unreduced gametes and establishment of successful polyploids often increase under unstable conditions (Ehrendorfer 1980; Marble 2004), especially under varying temperature (reviewed in Ramsey and Schemske 1998; Parisod, Holderegger et al. 2010). During the colonisation of the previously glaciated areas of Fennoscandia, the conditions were thus probably favourable for recurrent polyploidisation in *P. palustris*. Multiple independent polyploidisation events have been suggested for several other autopolyploid taxa, e.g. *Biscutella laevigata* L. (Parisod and Besnard 2007) and *Tolmiea menziesii* Torr. & A. Gray (Soltis and Soltis 1989), and are commonly considered the rule rather than the exception (Soltis and Soltis 1999; Soltis, Soltis et al. 2003).

The level of genetic diversity was similar in di- and tetraploid populations of *P. palustris*. A polyploidisation event followed by reproductive isolation from the diploid progenitors represents a severe bottleneck event, and subsequent loss of genetic diversity in the newly formed tetraploid is expected (Jakobsson, Hagenblad et al. 2006; Salmon and Ainouche 2010). Both recurrent polyploidisation events and gene flow from the diploid progenitors to the already existing tetraploid populations would counteract the bottleneck that a polyploidisation event represents (Soltis, Soltis et al. 2003; Husband 2004; Jørgensen, Ehrich et al. 2011). Triploids, which can be involved in both of these processes, are often found when di- and tetraploids occur in the same area (Husband 2004), and fertile triploids are reported from several autotetraploid taxa (reviewed in Soltis, Soltis et al. 2003). Studies of *Chamerion angustifolium* L. showed that triploids produce gametes with ploidal levels of *n*, 2*n* and 3*n* (Husband 2004), and thus may contribute to gene flow between the cytotypes. The existence of fertile triploids has in other studies been taken as support for hybridisation between ploidal levels (Henry, Dilkes et al. 2007; Slotte, Huang et al. 2008; Chapman and Abbott 2010). A two-step process involving a triploid "bridge" can also be involved in the formation of the tetraploid cytotype (reviewed in Soltis, Soltis et al. 2003; Soltis, Buggs et al. 2010). Husband (2004) found that triploid bridges contributed to a major part (62%) of autotetraploid

formation in *C. angustifolium*. He concluded that triploids could increase the rate of polyploid formation, even when the triploids are relatively unfit (Husband 2004). No experiments have explicitly confirmed the fertility of triploid *P. palustris*. However, in a series of crossing experiments, Hultgård (1987) reported that crossings between diploid and tetraploid *P. palustris* gave well developed seeds. Previous and current findings of triploids indicate that hybridisation between the cytotypes, and/or recurrent formation of the tetraploid cytotype, are likely processes in *P. palustris*. However, no proper criteria have been developed to distinguish between effects of recurrent origins of an autopolyploid and gene flow between the autopolyploid and its diploid progenitor (Soltis, Buggs et al. 2010).

Genetic admixture and high genetic diversity a result of secondary contact in Fennoscandia

The genetic variation within a species is often structured as a result of contemporary processes and historical events, such as cold periods and glaciations. Researchers agree that most of Fennoscandia was covered by ice from Middle Weichselian to the deglaciation at approximately 11,500 years ago (Lokrantz and Sohlenius 2006). Although there is evidence for ice-free areas within the maximum limits of the Late Weichselian ice sheet (i.e. the Last Glacial Maximum LGM), only very hardy organisms, if any, could have survived *in situ* in the North Atlantic area (e.g. Bennike 1999; Brochmann, Gabrielsen et al. 2003; but see Westergaard, Alsos et al. 2011). No ecological, cytological or distributional data suggest that *P. palustris* has been surviving the LGM in one of these ice-free areas (Hultgård 1987). The *P. palustris* populations in the Fennoscandian region display admixed genetic structure. It is likely that this genetic pattern reflects that the Fennoscandian populations have originated from two to three source regions, represented by the genetic homogenous clusters found in the East (Cluster1), the Alps (Cluster 2) and Great Britain (Cluster 3).

Large parts of Eastern Europe and Asia remained ice free during the last glaciation. Pollen analyses confirm the presence of *P. palustris* in arctic Russia 60,000 – 80,000 C^{14} years ago, i.e. in the period before the glaciation (Borgen and Hultgård 2003). Individuals from an eastern refugium could have immigrated into Northern Fennoscandia, resulting in the high degree of allocation to Cluster 1, which is otherwise dominating in the East region. Northern Russia has been an important source for immigration to the North Atlantic area also for other

subpolar/boreal plant species (Brochmann, Gabrielsen et al. 2003; Alsos, Eidesen et al. 2007) after the LGM.

Parnassia palustris could have immigrated from refugia in Central Europe into Southern Fennoscandia, resulting in the high degree of allocation to Cluster 2 in these regions. Several areas in the Alps are strong candidates as glacial refugia for alpine vascular plants (Tribsch and Schönswetter 2003) and have been suggested as source areas for many arctic-alpine species (Brochmann, Gabrielsen et al. 2003), e.g. *Vaccinium uliginosum* L. (Alsos, Engelskjøn et al. 2005) and *Ranunculus glacialis* L. (Schönswetter, Paun et al. 2003). In the current study, *P. palustris* populations sampled from the Alps comprised lower diversity and rarity than populations sampled from Fennoscandia. This was not expected if the Alps served as refugia during the LGM: Long-term isolated populations are expected to accumulate high numbers of rare genetic markers, while populations experiencing gene flow, or newly established populations, are expected to hold low such numbers (Schönswetter and Tribsch 2005). The low diversity and low degree of rarity of genetic markers found among the sampled Alp populations could reflect that these populations are the result of a recent immigration from other, more diverse source areas in the Alps or lowland Central Europe.

Pollen and seed records from the Middle and Late Weichselian, found in England, Eastern Scotland, the Netherlands, North-Western Germany and South-Western Denmark suggest a scenario where North-Western and Central Europe served as refugia for *P. palustris* during the last glaciation (References in Borgen and Hultgård 2003). The high degree of allocation to Cluster 3 (Great Britain) found in the South-Western Fennoscandian regions could be a result of immigration from a North-Western refugium. However, in their study of several alpine plant species from Scotland, Westergaard, Alsos et al. (2008) found no signs of *in situ* survival in this area during the LGM. They rather suggested a common glacial refugium in Central Europe for contemporary European/Eurasian and Scottish populations of alpine plants. A recent immigration to Scotland would explain the low genetic diversities found in Scottish populations of *P. palustris.* However, the current study does not provide results to neither support nor discard any of these hypotheses.

These suggestions are in concordance with conclusions based on isoenzyme markers (Borgen and Hultgård 2003), where two main immigration routes for *P. palustris* into Fennoscandia

were suggested: An eastern immigration route, supplying eastern parts of Fennoscandia (Sweden and Finland), and a southwestern main route over the land bridge connecting the British Isles and Southern Scandinavia to the continent, supplying Iceland and the western and north-western parts of Scandinavia.

Parnassia palustris is most abundant in the pioneer time span, decreasing when the vegetation approaches its maximum (Schwanck 1982; in Hultgård 1987). These ecological observations are supported by pollen records, suggesting that the postglacial colonisation of Fennoscandia is expected to have happened immediately after the retreat of the ice sheet (pollen records summarized in Table 20 in Hultgård 1987). The zone of overlap between the two cytotypes in Southern Scandinavia largely corresponds to the glacial boundary 11,500 years ago (Brochmann, Brysting et al. 2004). Based on the size of fossil pollen grains from South Central Sweden (Florin 1969) and on observation that the tetraploid cytotype is nearly absent from areas not covered by the ice during the LGM (Fig. 55 in Hultgård 1987), Hultgård inferred that the tetraploid cytotype of *P. palustris* is a result of *in situ* polyploidisation. During the deglaciation, the newly formed tetraploid cytotype must have been more successful than the diploid cytotype in colonizing the newly glaciated area, in order to create the current distribution pattern.

The genetic consequence of rapid colonisations of previously glaciated areas, like the one that probably occurred in *P. palustris*, is expected to be loss of genetic diversity, due to repeated bottlenecks (Leading-edge colonization; Hewitt 1995), especially for outcrossing diploids (Brochmann, Brysting et al. 2004). In spite of these expectations, the genetic diversity of *P. palustris* populations is actually higher in the previously glaciated areas in Fennoscandia than in the presumed source regions; Central Europe and East. High levels of genetic diversity is often found in areas that have served as refugia during the ice ages (Ehrich, Gaudeul et al. 2007), or in contact zones between populations that have been isolated in separate refugia (Abbott and Brochmann 2003; Petit, Aguinagalde et al. 2003; Parisod and Besnard 2007).The high levels of diversity found in Fennoscandian *P. palustris* populations have previously been explained as a result of maintenance of diversity during immigration from the South (Borgen and Hultgård 2003). The existence of a contact zone between plants immigrated from multiple refugia, e.g. in the South and the East, could perhaps better explain the high levels of diversity found in Fennoscandian populations in this and previous studies. Polyploidisation events

might have stabilized intraspecific hybrids between different diploid genetic lineages in the newly deglaciated areas, leading to new and favourable gene combinations. In *Vaccinium uliginosum* L., genetically admixed populations and enhanced genetic diversity was also taken as support for the existence of contact zones between genetic lineages from different refugia (Eidesen, Alsos et al. 2007).

According to Stebbins (1984), diploid populations could accumulate genome differences in isolated glacial refugia, causing new and favourable gene combinations when they meet and hybridize. When such intraspecific differentiated lineages meet in a contact zone, polyploidisation could help stabilize intraspecific hybrids with new and possibly adaptive gene combinations and higher levels of heterozygosity (Parisod and Besnard 2007; Parisod, Holderegger et al. 2010). The resulting higher heterozygosity might be a driving force for polyploids to expand into habitats that are unavailable for their diploid progenitors (Levin 2002). Higher levels of heterozygosity in autotetraploids than in their diploid progenitors are reported from other taxa (Soltis and Soltis 1989; Levin 2002). For instance, the level of heterozygosity in the autotetraploid *Tolmiea menziesii* Torr. & A. Gray was significantly higher than in the diploid cytotype of the same species (Soltis and Soltis 1989). Levels of heterozygosity was correlated with plant vigour in *Dactylis glomerata* L. (Tomekpe and Lumaret 1991). AFLP data do not provide information about the level of heterozygosity, but Hultgård and Borgen (2003) found greater isoenzyme heterozygosity in the European *P. palustris* tetraploids than in the diploids. The distribution of cytotypes, the admixed genetic structure and the higher levels of diversity and heterozygosity found in Fennoscandian populations of *P. palustris*, thus, all fits well with a hypothesis of multiple immigration lineages and admixture as a result of secondary contact.

Differences in survival between cytotypes caused by adaption to different climates?

The current study confirms that tetraploids are prevalent in the North-Western parts of Fennoscandia, and that di- and tetraploids have overlapping distribution in parts of Southern Norway and Sweden. The tetraploid populations from the Alps are found on the altitudinal limit of the species, in areas that were unglaciated or only partially glaciated during the LGM (Pers. Comm., Tribsch 2011; Schonswetter, Stehlik et al. 2005). No further conclusions can

be drawn on the distribution of cytotypes related to the boundaries of the Last Glacial Maximum.

The distribution of cytotypes in Fennoscandia lead Hultgård (1987) to infer an adaptive advantage of the tetraploid cytotype of *P. palustris* in colonizing previously glaciated areas. *In situ* ecological surveys of *P. palustris* in Fennoscandia showed that the tetraploid cytotype exhibited a wider ecological amplitude and less demands regarding soil composition than the diploid (Hultgård 1987). On the contrary, no such correlation between cytotype and habitat was found in an ecological study of Russian populations (Funamoto, Kondo et al. 2006), but only nine localities were included in this study. Based on ecological differences between the cytotypes, Hultgård (1987) proposed that the tetraploid cytotype of *P. palustris* was a result of recurrent intraspecific hybrid polyploidisation between diploid populations that had been geographically separated during the ice age. Because of greater isoenzyme heterozygosity in tetraploids than in diploids, Borgen and Hultgård (2003) hypothesized that due to greater genetic flexibility resulting from this heterozygosity, the intraspecific hybrid tetraploid populations would be able to respond more effectively to a changing environment than the diploids. The tetraploids would thus have an advantage during colonization of the previous glaciated land, causing the current distribution pattern. This is in line with the general, but nevertheless controversial, suggestion that polyploids have higher ecological tolerance and thus higher flexibility than their diploid progenitors (e.g.Soltis and Soltis 2000). In the present study, a null hypothesis (H_0) was formulated to test whether tetraploid *P. palustris* exhibits higher flexibility than diploid *P. palustris* towards varying temperatures or day lengths: There is no difference in survival and growth rate between diploid and tetraploid *P. palustris* when treated with different combinations of day length and temperature.

The *P. palustris* seedlings included in the growth experiment were growing extremely slow during the first months; the growth rate increased after a period of establishment. It is therefore not surprising that the variables measuring plant growth were positively correlated with initial size and age of the seedling. High temperature was positively affecting the leaf surface and dry weight, suggesting that temperature might be a limiting factor for *P. palustris*. Temperature increase is shown to have a positive effect on above-ground plant productivity in a number of arctic and temperate plant species (Rustad, Campbell et al. 2001; Rozema, Weijers et al. 2009). In the current study, ploidy did not affect any of the variables measuring

plant growth; no difference in growth rate between diploids and tetraploids during any of the treatments could be detected. The null hypothesis cannot be rejected according to the GLM analyses. However, the current experiments suffered from very low survival rate, and the GLM analyses were based on a data set consisting of only 42 individuals, which might be too little to detect an effect of ploidy.

The survival analyses included data from 222 plants, and thus represented an improvement of the statistical power compared to the GLMs. According the non-parametric log-rank approach, ploidy did not have an effect on survival. According to the more statistically powerful Cox' regression, however, several factors, including the interaction between temperature treatment and ploidy, significantly affected survival. Notably, long day treatment had a positive effect on survival, suggesting that day length might influence the distribution of *P. palustris*. It is known that long day treatment promote increased growth in plants that otherwise grow in short day conditions (Adams and Langton 2005). Overall, tetraploid plants showed a non-significant tendency towards higher survival than the diploids, while the tetraploids had significantly lower survival during the high temperature regime. In other studies, polyploidy is shown to change the temperature optima for several physiological processes (Levin 2002). In *Dactylis glomerata* L., the tetraploid cytotype is shown to have higher photosynthetic rate at lower temperatures compared to its diploid progenitor (Levin 2002). Several studies have shown that autotetraploids have increased resistance to frost compared to diploids (Chowdhury, Ghai et al. 1968) and low temperatures (Nishiyama 1944; Wit 1958) although also lower resistance to low temperatures has been reported for polyploids (Dvôrak and Fowler 1978; Tyler, Borrill et al. 1978). The tetraploids in the present study came from localities with lower temperature during the growth season than the diploids. If tetraploid populations of *P. palustris* are better equipped to handle changing temperatures than the diploids, one would expect the tetraploids to have as high or higher survival than the diploids under different treatments, also under treatments that they are not adapted to (i.e. high temperature). The higher survival of diploids during the high temperature treatment, and the higher survival of tetraploids during the low temperature treatment, are consistent with temperature differences between the areas from which the populations are collected. The differences in survival could very well be merely an effect of adaptions to different climates, and cannot be assigned to ploidy *per se*. The null hypothesis can, thus, not be rejected based on the Cox' regression analyses.

Due to poor seed germination and seedling survival, the experimental setup presented in the current study turned out not to be optimal. Furthermore, it is possible that the temperatures and day lengths used in the present study were unsuitable to detect higher ecological flexibility in tetraploids, or that the cytotypes differ in response to other factors or combination of factors than the ones tested. Other experimental studies have indeed indicated that the polyploids might differ from their progenitors with respect to other factors than day length and temperature (Levin 2002). In a series of transplant experiments, Ramsey (2011) showed that genome duplication *per se* contributed to increased survival in autopolyploid *Achillea borealis* Bong. in dune habitats. In a controlled growth study of *Bromopsis erecta* (Huds.) Fourr. (syn. *Bromus erectus* Huds.), however, no effect of ploidy was found on plant response to competition (Münzbergová 2006). Several studies indicate that some polyploids are more tolerant to drought than their diploid progenitors; however, both positive and negative correlations between ploidy and nutrient tolerance have been reported (Levin 2002). The evidence, which so far have been used to discuss a possible adaptive advantage of polyploids under certain conditions, are mainly anecdotal (te Beest, Le Roux et al. 2011) and the few experimental studies that have so far been conducted, including the present one, are either not conclusive or provides contradiction evidence. Many more well-designed experiments are needed before we can either confirm or reject the possibility that polyploidy *per se* may provide an adaptive advantage in a changing climate.

Conclusions

Species rank for autotetraploid *P. palustris* cannot be justified, and only one species is found in the study area. Tetraploid populations were found in areas from which only diploid *P.palustris* has otherwise been reported but the overall distribution pattern of the two cytotypes corresponded with previous studies. The di- and tetraploid populations of *P. palustris* have overall similar levels of genetic diversity, and the tetraploid populations from the Alps and Scandinavia, respectively, are more similar to sympatric diploid populations in genetic structure than they are to each other. Recurrent polyploidisations and/or hybridisation between the cytotypes might be causing these patterns. Tri- and pentaploid individuals found in sympatric areas support both of these scenarios. Different immigration routes have lead to the mixed genetic structure found in Fennoscandia. The existence of a contact zone between immigrants from southern and eastern refugia could explain the high levels of diversity found in Fennoscandian populations. Higher genetic diversity caused by secondary contact could in theory cause higher ecological flexibility in tetraploids and explain the current distribution pattern. However, no effect of ploidy *per se* on flexibility in response to varying day lengths and temperatures could be detected in tetraploid *P. palustris*.

Literature cited

Abbott, R. J. and C. Brochmann (2003). "History and evolution of the arctic flora: in the footsteps of Eric Hultén." Molecular Ecology 12: 299 - 313.

Adams, K. L. and J. F. Wendel (2005). "Polyploidy and genome evolution in plants." Current Opinion in Plant Biology 8: 135-141.

Adams, S. R. and F. A. Langton (2005). "Photoperiod and plant growth: a review." Journal of horticultural science & biotechnology 80: 2 - 10.

Akaike, H. (1974). "A new look at the statistical model identification." IEEE Transactions in Automatic Control 19: 716 - 723.

Alsos, I. G., P. B. Eidesen, et al. (2007). "Frequent long-distance plant colonization in the changing Arctic." Science 316: 1606 - 1608.

Alsos, I. G., T. Engelskjøn, et al. (2005). "Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species." Molecular Ecology 14: 2739-2753.

Baack, E. J. and M. L. Stanton (2005). "Ecological factors influencing tetraploid speciation in Snow Buttercups (Ranunculus adoneus): Niche differentiation and tetraploid establishment." Evolution 59: 1936 - 1944.

Barlow, R. E., A. W. Marshall, et al. (1963). "Properties of probability distributions with monotone hazard rate." The Annals of Mathematical Statistics 34: 375 - 389.

Bendiksby, M., A. Tribsch, et al. (2011). "Allopolyploid origin of Galeopsis tetraploids – revisiting Muntzing's (1932) classical textbook example using molecular tools." New Phytologist 191: 1150 - 1167.

Bennike, O. (1999). "Colonisation of Greenland by plants and animals after the last ice age: a review

" Polar Record 35: 323 - 336.

Bonin, A., E. Bellemain, et al. (2004). "How to track and assess genotyping errors in population genetics studies." Molecular Ecology 13: 3261 - 3273.

Borgen, L. and U.-M. Hultgård (2003). "Parnassia palustris: A genetically diverse species in Scandinavia." Botanical Journal of the Linnean Society 142: 347 - 372.

Bretagnolle, F. and R. Lumaret (1995). "Bilateral polyploidization in Dactylis glomerata L. subspecies lusitanica: Occurrence, morphological and genetic characteristics of first polyploids." Euphytica 84: 197 - 207.

Brochmann, C., A. K. Brysting, et al. (2004). "Polyploidy in arctic plants." Biological Journal of the Linnean Society 82: 521 - 536.

Brochmann, C., T. M. Gabrielsen, et al. (2003). "Glacial survival or tabula rasa? The history of North Atlantic biota revisited." Taxon 52: 417 - 450.

Chapman, M. A. and R. J. Abbott (2010). "Introgression of fitness genes across a ploidy barrier." New Phytologist 186: 63-71.

Chowdhury, J. B., B. S. Ghai, et al. (1968). "Studies on the cytology and fertility in the induced polyploids of self-incompatible Brassica campestris var. Brown sarson." Cytologia 33: 269 - 275.

Corander, J. and P. Marttinen (2006). "Bayesian identification of admixture events using multilocus molecular markers." Molecular Ecology 15: 2833 - 2843.

Corander, J., P. Marttinen, et al. (2008). "Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations." BioMed Central Bioinformatics 9: 539 - 544.

Cox, D. R. (1972). "Regression models and life-tables (with discussion)." Journal of the Royal Statistical Society. Series: B (Methodological). 34: 187 - 220.

Crawley, M. J. (2007). The R Book. Chichester, John Wiley & Sond Ltd.

Dalgaard, P. (2008). Introductory statistics with R. New York, Springer.

Darlington, C. D. (1932). Recent advances in Cytology. Philadelphia, P. Blakiston's son & co. inc.

Dice, L. R. (1945). "Measures of the amount of ecologic association between species." Ecology 26: 297 - 302.

Dijk, P. V. and T. B. Schotsman (1997). "Chloroplast DNA phylogeography and cytotype geography in autopolyploid Plantago media." Molecular Ecology 6: 345 - 352.

Dobes, C. and B. Hahn (1997). "IOPB chromosome data 11." International Organization of Plant Biosystematists Newsletter 26/27s: 15 - 18.

Doyle, J. J. and J. L. Doyle (1987). "A rapid DNA isolation procedure for small quantities of fresh leaf tissue." Phytochemical Bulletin 19: 11 - 15.

Dvôrak, J. and D. B. Fowler (1978). "Cold hardiness potential of triticale and tetraploid rye." Crop Science 17: 477 - 478.

Ehrendorfer, F. (1980). Polyploidy and distribution. New York, Plenum Press.

Ehrich, D. (2006). "AFLPdat: A collection of R functions for convenient handling of AFLP data." Molecular Ecology Notes: 603 - 604.

Ehrich, D., M. Gaudeul, et al. (2007). "Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains." Molecular Ecology: 2542-2559.

Eidesen, P. B., I. G. Alsos, et al. (2007). "Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species." Molecular Ecology 16: 3902 - 3925.

Elven, R. (2007 onwards, 9. November 2011). "Checklist of the panarctic flora (PAF) Vascular plants [WWW document] " Elven, R. Retrieved 15 November, 2011, from http://www.binran.ru/infsys/paflist/index.htm

Erlandsson, S. (1942). "Cytologiskt-växtgeografiska rasstudier i Nordens Parnassia palustris-population." Acta Horticulturae Bergiani 13: 117 - 148.

Evanno, G., S. Regnaut, et al. (2005). "Detecting the number of clusters of individuals using the software Structure: A simulation study." Molecular Ecology 14: 2611 - 2620.

Excoffier, L. (2007). Analysis of population subdivision. Chichester, John Wiley & sons Ltd.

Excoffier, L., P. Smouse, et al. (1992). "Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. ." Genetics 131: 479 - 491.

Favarger (1967). "Cytologie et distribution des plantes." Biological review 42: 163 - 206.

Fawcett, J. A., S. Maere, et al. (2009). "Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event." Proceedings of the National Academy of Sciences 106: 5737 - 5742.

Field, A. (2005). Discovering statistics using SPSS (and sex, drugs and rock'n roll). London, SAGE Publications.

Florin, M. B. (1969). "Late-glacial and pre-boreal vegetation in Central Sweden. I. Records of Pollen Species." Svensk Botanisk Tidsskrift 63: 143 - 187.

Flovik, K. (1940). "Chromosome numbers and polyploidy within the flora of Spitzbergen." Hereditas 26: 430 - 440.

Funamoto, T., K. Kondo, et al. (2006). "Intraspecific polyploidy of Parnassia palustris var. multiseta (Saxifragaceae s.l.) collected in Primorye and Altai territories, Russia." Chromosome Botany 1: 23 - 26.

Gadella, T. W. J. and E. Kliphuis (1968). "Parnassia palustris in the Netherlands." Acta Botanica Neerlandica 17: 165 - 172.

Gates, R. R. (1909). "The stature and chromosomes of Oenothera gigas Dr Vries." 3.

Gornall, R. J. and J. E. Wenthworth (1992). "Variation in the chromosome number of Parnassia palustris L. in the British Isles." New Phytologist 123: 383 - 388.

Grant, V. (1981). Plant speciation. New York, Colombia University Press.

Hagerup, O. (1931). "Über Polyploidie in Beziehung zu Klima, Ökologie und Phylogenie." Hereditas 16: 19 - 40.

Hammer, Ø., D. A. T. Harper, et al. (2001). "PAST: Paleontological statistics software package for education and data analysis." Palaeontologia Electronica 4.

Hendry, A. P. (2009). "Evolutionary biology: Speciation." Nature 458: 162 - 164.

Henry, I. M., B. P. Dilkes, et al. (2007). "Genetic basis for dosage sensitivity in Arabidopsis thaliana." Public Library of Science Genetics 3: 593 - 602.

Hewitt, G. M. (1995). "Some genetic consequences of ice ages, and their role in divergence and speciation." Biological Journal of the Linnean Society 58: 247 - 276.

Holland, B., A. Clarke, et al. (2008). "Optimizing automated AFLP scoring parameters to improve phylogenetic resolution." Systematic Biology 57: 347 - 366.

Hultén, E. (1971). "The cirkumpolar plants, II." Kongliga Svenska Vetenskaps-akademiens Handlingar 13.

Hultgård, U.-M. (1987). Parnassia palustris L. in Scandinavia. Department of Systematic Botany. Uppsala, Acta Universitatis Upsaliensis. PhD thesis: 128.

Husband, B. C. (2004). "The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations." Biological Journal of the Linnean Society 82: 537 - 546.

Huson, D. H. (1998). "Splitstree: A program for analyzing and visualizing evolutionary data." Bioinformatics 14: 68 - 73.

Huson, D. H. (2005). "Application of phylogenetic networks in evolutionary studies." Molecular Biology and Evolution 23: 254 - 267.

Huson, D. H. and D. Bryant (2006). "Application of phylogenetic networks in evolutionary studies." Molecular Biology and Evolution 23: 254 - 267.

Jakobsson, M., J. Hagenblad, et al. (2006). "A unique recent origin of the allotetraploid species Arabidopsis suecica: Evidence from nuclear DNA markers." Molecular Biology and Evolution 23: 1217-1231.

Johnson, A. W. and J. G. Packer (1965). "Polyploidy and environment in arctic Alaska." Science 148: 237 - 239.

Jørgensen, M., D. Ehrich, et al. (2011). "Interspecific and interploidal gene flow in Central European Arabidopsis (Brassicaceae)." BMC Evolutionary Biology (in press).

Jørgensen, M. H., R. Elven, et al. (2006). "Taxonomy and evolutionary relationships in the Saxifraga rivularis complex." Systematic Botany 31: 702 - 729.

Kosman, E. and K. J. Leonard (2005). "Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species." Molecular Ecology 14: 415 - 424.

Laval, L. G. E. and S. Schneider (2005). "Arlequin ver. 3.0: An integrated software package for population genetics data analysis." Evolutionary Bioinformatics Online 1: 47 - 50.

Levin, D. A. (1983). "Polyploidy and novelty in flowering plants." The Americal Naturalist 122: 1 - 25.

Levin, D. A. (2002). The role of chromosomal change in plant evolution. New York, Oxford University Press.

Li-Bing, Z. and M. P. Simmons (2006). "Phylogeny and delimitation of the Celastrales inferred from nuclear and plastid genes." Systematic Botany 31: 122 - 137.

Lokrantz, H. and G. Sohlenius (2006). Ice marginal fluctuations during the Weichselian glaciation in Fennoscandia, a literature review. Technical Report. G. S. o. S. (SGU). Stockholm, SGU.

Lutz, A. (1907). "A preliminary note on the chromosomes of Oenothera Lamarckiana and one of its mutants, O. Gigas." Science 26: 151 - 152.

Lyrene, P. M., N. Vorsa, et al. (2003). "Polyploidy and sexual polyploidization in the genus Vaccinium." Euphytica 133: 27 - 36.

Löve, A. (1950). "Some innovations and nomenclatural suggestions in the Icelandic flora." Botaniska Notiser: 24 - 60.

Löve, A. and D. Löve (1943). "The significance of differences in distribution of diploids and polyploids." Hereditas 29: 145 - 163.

Mallet, J. (2007). "Hybrid speciation." Nature 446: 279-283.

Manton, I. (1950). Problems of cytology and evolution in the Pteridophyta. Cambridge, Cambridge University Press.

Marble, B. K. (2004). "'Why polyploidy is rarer in animals than in plants': myths and mechanisms." Biological Journal of the Linnean Society 82: 453 - 466.

Mariette, S., V. L. Corre, et al. (2002). "Sampling within the genome for measuring within population diversity." Molecular Ecology 11: 1145 - 1156.

Mayr, E. (1942). Systematics and the origin of species, from the viewpoint of a zoologist. New York, Coloumbia University Press.

Meirmans, P. G., J. Goudet, et al. (2011). "Ecology and life history affect different aspects of the population structure of 27 high-alpine plants." Molecular Ecology 20: 3144-3155.

Moran, J. L., P. J. Solomon, et al. (2007). "New models for old questions: generalized linear models for cost prediction." Journal of Evaluation in Clinical Practice 13: 381-389.

Mueller, U. G. and L. L. Wolfenbarger (1999). "AFLP genotyping and fingerprinting." TREE 14: 389 - 394.

Murray, M. G. and W. F. Thompson (1980). "Rapid isolation of high molecular weight plant DNA." Nucleic Acids Research 8: 4321 - 4325.

Müntzing, A. (1936). "The evolutionary significance of polyploidy." Hereditas 21: 363 - 378.

Münzbergová, Z. (2006). "No effect of ploidy level in plant response to competition in a common garden experiment." Biological Journal of the Linnean Society 92: 211 - 219.

Nei, M. (1987). Molecular evolutionary genetics. New York, Colombia University Press.

Nishiyama, I. (1944). "Study on artificial polyploid plants VI: On the different growth of the diploid and tetraploid radish in the winter season." Journal of the Japanese Society for Horticultural Science 13: 245 - 252.

Parisod, C. and G. Besnard (2007). "Glacial in situ survival in the Western Alps and polytopic autopolyploidy in Biscutella laevigata L. (Brassicaceae)." Molecular Ecology 16: 2755-2767.

Parisod, C., R. Holderegger, et al. (2010). "Evolutionary consequences of autopolyploidy." New Phytologist 186: 5-17.

Petit, R. J., I. Aguinagalde, et al. (2003). "Glacial refugia: Hotspots but not melting pots of genetic diversity." Science 300: 1563 - 1565.

Pires, J. C., J. Zhao, et al. (2004). "Flowering time divergence and genomic rearrangements in resynthesized Brassica polyploids (Brassicaceae)." Biological Journal of the Linnean Society 82: 675 - 688.

Pritchard, J. K., M. Stephens, et al. (2000). "Inference of population structure using multilocus genotype data." Genetics 155: 945 - 959.

Ramsey, J. (2011). "Polyploidy and ecological adaptation in wild yarrow." Proceedings of the National Academy of Sciences 108: 7096-7101.

Ramsey, J. and D. W. Schemske (1998). "Pathways, mechanisms, and rates of polyploid formation in flowering plants." Annual Reviews of Ecology and Systematics 29: 467 - 501.

Ramsey, J. and D. W. Schemske (2002). "Neopolyploidy in Flowering Plants." Annual Review of Ecology and Systematics 33: 589-639.

Rosenberg, N. A. (2002). "Genetic structure of human populations." Science 298: 2381 - 2385.

Rozema, J., S. Weijers, et al. (2009). "Annual growth of Cassiope tetragona as a proxy for Arctic climate: developing correlative and experimental transfer functions to reconstruct past summer temperature on a millennial time scale." Global Change Biology 15: 1703-1715.

Rustad, L., J. Campbell, et al. (2001). "A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming." Oecologia 126: 543-562.

Saghai-Maroof, M. A., K. M. Soliman, et al. (1984). "Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics." Population Biology 81: 8014 - 8018.

Salmon, A. and M. L. Ainouche (2010). "Polyploidy and DNA methylation: new tools available." Molecular Ecology 19: 213 - 215.

Schonswetter, P., I. Stehlik, et al. (2005). "Molecular evidence for glacial refugia of mountain plants in the European Alps." Molecular Ecology 14: 3547 - 3555.

Schwanck, B. (1982). Strändernas floradynamik på flacka kuster med maximal landhöjning. Landhöjning och kustbygdsförändring. Nordiskt symposium. Luleå, Sweden. 1.

Schwarz, G. (1978). "Estimating the dimensions of a model." Annals of Statistics 6: 461 - 464.

Schönswetter, P., O. Paun, et al. (2003). "Out of the Alps: colonization of Northern Europe by East Alpine populations of the Glacier Buttercup Ranunculus glacialis L. (Ranunculaceae)." Molecular Ecology 12: 3373 - 3381.

Schönswetter, P., J. Suda, et al. (2007). "Circumpolar phylogeography of Juncus biglumis (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers." Molecular Phylogenetics and Evolution 42: 92 - 103.

Schönswetter, P. and A. Tribsch (2005). "Vicariance and dispersal in the alpine perennial Bupleurum stellatum L.(Apiaceae)." Taxon 54: 725 - 732.

Slotte, T., H. Huang, et al. (2008). "Polyploid speciation did not confer instant reproductive isolation in Capsella (Brassicaceae)." Molecular Biology and Evolution 25: 1472 - 1481.

Soltis, D. E., V. A. Albert, et al. (2009). "Polyploidy and angiosperm diversification." American Journal of Botany 96: 336 - 348.

Soltis, D. E., R. J. A. Buggs, et al. (2010). "What we still don't know about polyploidy." Taxon 59: 1387 - 1403.

Soltis, D. E. and P. S. Soltis (1989). "Genetic consequences of autopolyploidy in Tolmiea (Saxifragaceae)." Evolution 43: 586 - 594.

Soltis, D. E. and P. S. Soltis (1999). "Polyploidy: Recurrent formation and genome evolution." Trends in Ecology and Evolution 14: 348 - 352.

Soltis, D. E., P. S. Soltis, et al. (2007). "Autopolyploidy in angiosperms: Have we grossly underestimated the number of species?" Taxon 56: 13 - 30.

Soltis, D. E., P. S. Soltis, et al. (2003). "Advances in the study of polyploidy since Plant Speciation." New Phytologist 161: 173-191.

Soltis, P. S. and D. E. Soltis (2000). "The role of genetic and genomic attributes in the success of polyploids." Proceedings of the National Academy of Sciences 97: 7051 - 7057.

Stebbins, G. L. J. (1947). "Types of polyploids: Their classification and significance." Advances in Genetics 1: 403 - 429.

Stebbins, G. L. J. (1949). "The evolutionary significance of natural and artificial polyploids in the family Gramineae." Hereditas 35: 461 - 485.

Stebbins, G. L. J. (1971). Chromosomal evolution in higher plants. London, Addison - Wesley.

Stebbins, G. L. J. (1984). "Polyploidy and the distribution of arctic-alpine flora - new evidence and a new approach." Botanica Helvetica 94: 1 - 13.

Stebbins, G. L. J. (1985). "Polyploidy, hybridization, and the invasion of new habitats." Annals of the Missouri Botanical Garden 72: 824 - 832.

Stehlik, I., J. J. Schneller, et al. (2002). "Immigration and in situ glacial survival of the low-alpine Erinus alpinus (Scrophulariaceae)." Biological Journal of the Linnean Society 77: 87 - 103.

te Beest, M., J. J. Le Roux, et al. (2011). "The more the better? The role of polyploidy in facilitating plant invasions." Annals of Botany: 1 - 27.

Thompson, J. N. and R. Lumaret (1992). "The evolutionary dynamics of polyploid plants: Origins, establishment and persistence." Trends in Ecology and Evolution 7: 302 - 307.

Tomekpe, K. and R. Lumaret (1991). "Association between quantitative traits and allozyme heterozygosity in a tetrasomic species: Dactylis glomerata." Evolution 45: 359 - 370.

Tribsch, A. and P. Schönswetter (2003). "In search for Pleistocene refugia for mountain plants: patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence in the Eastern European Alps." Taxon 52: 477 - 497.

Tyler, B., M. Borrill, et al. (1978). "Studies in Festuca. X. Observations on germination and seeling cold tolerance in diploid fesctuca pratensis and tetraploid F. pratensis var. appennina in relation to their altitudinal distribution." Journal of Applied Ecology 15: 219 - 226.

Wahlenberg, G. (1812). Flora Lapponica. Stockholm, Berolini.

Westergaard, K. B., I. G. Alsos, et al. (2008). "Genetic diversity and distinctiveness in Scottish alpine plants." Plant Ecology & Diversity 1: 329-338.

Westergaard, K. B., I. G. Alsos, et al. (2011). "Glacial survival may matter after all: nunatak signatures in the rare European populations of two west-arctic species." Molecular Ecology 20: 376 - 393.

Westergaard, K. B., M. H. Jørgensen, et al. (2010). "The extreme Beringian/Atlantic disjunction in Saxifraga rivularis (Saxifragaceae) has formed at least twice." Journal of Biogeography 37: 1262 - 1276.

Wheeler, Q. D. and R. Meier (2000). Species concepts and phylogenetic theory: A debate. New York, Colombia University Press.

Wilson, K. and I. C. W. Hardy (2002). Statistical analyses of sex ratios: an introduction. Cambridge, Cambridge University Press.

Wit, F. (1958). "Tetraploid Italian ryegrass (Lolium multiflorum Lam.)." Euphytica 7: 47 - 58.

Wu, D., H. Wang, et al. (2005). "Pollen morphology of Parnassia L. (Parnassiaceae) and its systematic implications." Journal of Integrative Plant Biology 47: 2 - 12.

Ziegenhagen, B., P. Guillemaut, et al. (1993). "A procedure for mini-preparations of genomic DNA from needles of silver fir (Abies alba Mill.)." Plant Molecular Biology Reporter 11.

Appendix

Table of contents

Table S1. *Parnassia palustris* individuals used for estimation of ploidy and/or AFLP analyses. Ploidy estimations were carried out in several rounds, using fresh and silica dried material. Flow cytometry using DAPI staining on fresh material was performed in November 2009, in April 2010 and in April 2011. Flow cytometry using DAPI staining on silica dried material was performed in April 2011. Based on the original allocation of individuals, populations were assigned to one of the four Structure clusters in the following way: The allocation to each cluster for all the individuals in a population was summed up. The total allocation to one cluster for all the individuals was then divided by the total allocation to all four clusters, resulting in the population's fraction of allocation to this one cluster. When this fraction was higher than 60%, the population was assigned to this cluster. Populations that could not be assigned to any cluster according to this criterion were set as mixed. N/A - not available: populations were not assigned to a Structure group either because reliable AFLP data could not be obtained, or because they were not included in the Structure analysis (Japan and Alaska).

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
1							Mixed
	1_1		4x				
	1_2	1		4x			
	1_3	2	4x				
	1_4	1	4x				
	1_6		4x				
	1_7		4x				
	1_8		4x				
	1_9		4x				
	1_11		4x				
2							Mixed
	2_1			4x			
	2_2	1		4x			
	2_5	1		4x	4x		
	2_7				4x		
	2_8		4x	4x			
	2_9		2x	4x			
	2_10	1	4x				
	2_11		4x				
3							2
	3_1		4x				
	3_2			4x			
	3_3		4x				
	3_4		4x				
	3_5		4x				
	3_6		4x				
	3_7		4x				
	3_8		4x				
	3_9		4x				

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	3_10		4x				
	3_11		4x				
	3_12		4x				
	3_13		4x				
	3A	1		4x			
	3B			4x			
	3C			4x			
4							Mixed
	4_1		4x				
	4_2			4x			
	4_3		4x				
	4_4		4x		4x		
	4_5		4x		4x		
	4_6		4x				
	4_7	1					
	4_8		4x				
	4_9	1	4x		4x		
	4_10		4x		4x		
	4_11		4x				
	4A			4x			
	4B			4x			
	4C			4x			
	4D			4x			
	4E			4x			
	4F			4x			
	4G	1		4x			
	4H			4x			
	4I			4x			
	4J			4x			
	4K			4x			
	4L			4x			
	4M			4x			
	4N			4x			
	4O			4x			
	4P			4x			
	4Q			4x			
5							2
	5_1	1	4x				
	5_2	1		4x			
7							Mixed
	7_1		4x				

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	7_2	1		4x			
	7_3	1	4x				
	7_4		4x				
	7_5	1	4x				
8							Mixed
	8_1		4x		4x		
	8_2			4x			
	8_3		4x				
	8_4		4x				
	8_5		4x		4x		
	8_6	1	4x				
	8_7	1	4x		4x		
	8_8		4x		4x		
	8_9		4x		4x		
	8_10		4x				
	8_11		4x				
	8_12		4x				
	8_13		4x				
9							Mixed
	9_1	1	4x				
10							2
	10_1	1	2x				
	10_2	1		2x			
	10_3	1	2x				
	10_4	1	2x				
	10_5	1	2x				
	10_6		2x				
	10_7		2x				
	10_8		2x				
	10_9		2x				
	10_10		4x				
11							2
	11_1	1	2x				
	11_2	2		2x			
	11_3	1	2x				
	11_4	1	2x				
12							Mixed
	12_1		4x				
	12_2			4x			
	12_3		4x				
	12_4		4x				
	12_6	1	4x				
	12_10		4x				

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	12_12	1					
	12A			4x			
	12B			4x			
	12C			4x			
	12D			4x			
	12E			4x			
	12F			4x			
	12H			4x			
	12I			4x			
	12K	1		4x			
	12_11B			4x			
	12_11C			4x			
13							2
	13_1				2x		
	13_2			2x	2x		
	13_3		2x		2x		
	13_4				2x		
	13_5		2x				
	13_6		2x				
	13_7		2x				
	13_8		2x				
	13_9	1	2x				
14							2
	14_1		2x		2x		
	14_2	1		2x			
	14_3		2x				
	14_4		2x				
	14_5		2x				
	14_6		2x				
	14_7		2x				
	14_8		2x				
	14_10		2x				
15							1
	15_1	1			4x		
	15_2	1		4x			
	15_3	1			4x		
16							1
	16_1	1					
	16_4	1					
	16_7	1					
	16A			4x			
	16B			4x			
	16C	1		4x			

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	16D			4x			
	16E			4x			
	16F			4x			
	16G			4x			
	16H			4x			
	16I			4x			
	16J			4x			
	16K			4x			
	16L			4x			
	16M			4x			
	16N			4x			
	16T	1					
17							1
	17_3	1					
	17_4	1					
	17_5	1					
	17B			4x			
	17C			4x			
	17D			4x			
	17E			4x			
	17F	1		4x			
	17G			4x			
	17H			4x			
	17I			4x			
	17J			4x			
	17K			4x			
	17L	1		4x			
	17M			4x			
18							1
	18_1	1					
	18_1A	1		4x			
	18_4	1					
	18_5		4x				
	18_8	1					
	18A			4x			
	18B			4x			
	18C			4x			
	18D			4x			
	18E			4x			
	18F			4x			
	18I			4x			
19							4
	19_1			1			

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	19B	1		4x			
20							1
	20_1	1	4x				
	20_2	1		4x			
	20_3	1	4x				
	20_5	1					
	20_6	1					
	20A			4x			
	20_2A	1					
21							3
	21_1	1	4x				
22							1
	22_2	1	4x				
	22A	1		4x			
	22B			4x			
23							1
	23A_1	1	4x				
	23B_5	1					
	23B_6	1					
	23B_A			4x			
	23B_A2	1					
	23B_B			4x			
	23B_C	1		4x			
	23B_D			4x			
	23B_E			4x			
	23B_F			4x			
	23B_G			4x			
	23B_H			4x			
	23B_I			4x			
	23C_1	1					
26							4
	26_5	1					
	26J			4x			
	26K			4x			
	26L			4x			
	26M			4x			
	26N			4x			
	26O			4x			
	26P			4x			
	26Q			4x			
	26A			4x			
	26B	1		4x			
	26.6A			4x			

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	26.6C			4x			
	26.10A			4x			
	26.10B			4x			
	26.10C			4x			
	26.10D			4x			
	26.10E			4x			
	26.10F			4x			
	26.10G			4x			
	26.10H			4x			
	26.10I	1		4x			
27							4
	27_1	1					
	27A			4x			
	27B	1		4x			
	27C	1		4x			
	27D	1		4x			
28							4
	28_4		4x		4x		
	28_5	1		4x	4x		
29							N/A
	29_2				4x		
30							Mixed
	30_1r	1					
	30.1A			4x			
	30.1C			4x			
	30.1D			4x			
	30.1E			4x			
	30.1F			4x			
	30.1G			4x			
	30.1H			4x			
31							
	31A			4x	4x		
	31B			4x	4x		
	31C			4x			
	31D			4x			
	31E			4x	4x		
	31F			4x	4x		
	31.1B			4x			
32							Mixed
	32A_1				2x		
	32A_4			2x			
	32A_5				2x		
	32B_1	2					

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	32B_2	1			>5x		
	32C_1	1			2x		
	32C_2	2					
36							N/A
	36A			4x			
37							N/A
	37A	1					
	37G			2x			
	37B			2x			
	37C	1		2x			
	37D			2x			
	37E	1		2x			
	37F	1		2x			
	37H			2x			
	37I			2x			
	37J			2x			
	37K			2x			
	37L			2x			
	37M			2x			
	37N			2x			
	37O			2x			
	37P			2x			
	37Q			2x			
	37R	1		2x			
38							2
	38_2				2x		
	38_3			2x			
	38_4				2x		
	38A	1		2x			
	38B	1		2x			
	38C			2x			
	38D			2x			
	38E			2x			
	38P			2x			
	38Q			2x			
39							N/A
	39A			2x			
	39B			2x			
	39C			2x			
	39D			2x			
	39E			2x			
41							3
	41_1	1					

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	41_4	1		2x			
	41_5	1			2x		
	41_6				2x		
	41_7				2x		
	41_8				2x		
42							3
	42_5	1			2x		
	42_6	1		2x			
	42_7				2x		
	42_8				2x		
	42_9				2x		
	42_10	1					
43							Mixed
	43_1				3x		
	43_2	1		2x			
	43_3				2x		
	43_4				2x		
	43_6				2x		
	43_9	1					
	43_10	1					
44							Mixed
	44A			4x			
	44B			4x			
	44C			4x			
	44D	1		4x			
	44E	1		4x			
	44F	1		4x			
	44G			4x			
	44H			4x			
	44I			4x			
	44J			4x	4x		
	44K			4x			
	44L	1		4x	4x		
	44M			4x	4x		
	44N			4x			
	44O			4x	4x		
	44P			4x	4x		
	44Q			4x			
46							2
	46_1	1					
	46_3	1					
	46_4	1					
	46_4				2x		

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	46_5	1			2x		
	46_8				2x		
	46_10				2x		
	46_11				2x		
47							2
	47_3	1			2x		
	47_4	1			2x		
	47_5				2x		
	47_10	1					
48							2
	48_1	1					
	48_2				2x		
	48_3	1					
	48_5	1					
	48_7	1					
	48_8				2x		
	48_9				2x		
49							2
	49_1				4x		
	49_4	1					
	49_5	1					
	49_7	1			4x	4x	
	49_8				4x		
	49_9				4x		
	49_10	1			4x		
50							2
	50_1				4x		
	50_2				4x		
	50_3	1			4x		
	50_4				4x		
	50_5				5x		
	50_9	1					
	50_10	1					
	50_11	1					
51							1
	51A_1	1			4x		
	51A_2	1			4x		
	51A_3	1					
	51A_5_1	1			4x		
	51A_5_2				4x		
	51A_5_3				4x		
	51A_5_4				4x		
	51B_2	1					

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	51B_5	1					
52							Mixed
	52_1				4x		
	52_2	1					
	52_3	1			4x		
	52_4	1			4x		
	52_5	1			4x		
	52_6	1			4x		
53							Mixed
	53_1	1			4x		
	53_2				4x		
	53_3	1			4x		
54							Mixed
	54_1	1			4x		
	54_2	1			4x		
	54_3	1			4x		
55							2
	55_1				2x		
	55_2	1					
	55_3	1			2x		
	55_4	1			2x		
	55_5	1			2x		
	55_6	1					
	55_7	1			2x		
57							3
	57_1				2x		
	57_2	1			2x		
	57_3	1			2x		
	57_4	1			2x		
	57_5	1			2x		
58							N/A
	58_1				2x		
	58_2				2x		
	58_3				2x		
	58_4				2x		
	58_5				2x		
59							3
	59_1				2x		
	59_2	1			2x		
	59_3	1			2x		
	59_4	1					
	59_5	1			2x		

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
60							1
	60_7	1			4x		
	60_11				4x		
	60_12				4x		
	60_17						
61							3
	61_2				4x		
	61_3	1					
	61_4	1			4x		
62							3
	62_5	1			4x		
	62_6	1			4x		
	62_7	1			4x	4x	
	62_8				4x	4x	
	62_9				4x		
	62_10	1			4x		
65							N/A
	65_1				4x		
	65_3				4x		
	65_4				4x		
	65_5				4x		
66							3
	66_1	1			2x		
	66_2				2x		
	66_3				2x		
	66_4				2x	2x	
	66_5				3x		
	66_6				3x	3x	
	66_7				2x		
	66_8	1			2x		
	66_9				2x		
	66_10	1			2x		
	66_11				2x		
	66_12				2x		
	67_1	1					
67							N/A
	67_3				2x		
	67_4	1			2x		
	67_5				2x		
68							3
	68_1a				4x		
	68_1	1					

Pop	Sample ID	AFLP *	Fresh 2009	Fresh 2010	Silica 2011	Fresh 2011	Structure cluster
	68_2	1			4x		N/A
70							
	70_1				2x		

* 1 - Individuals scored for 144 markers, 2 - Individuals scored for 126 markers.

Table S2. List of collectors involved in sampling of seed, silica dried material and living plants of *Parnassia palustris*

Andreas Hilpold: AH

Andreas Tribsch: AT

Anna Ronikier: AR

Anne Krag Brysting: AKB

Astrid Bjørgaas: AB

Aveliina Helm: AH

Barbara Mable: BM

Christian E. Pettersen: CEP

David Murray: DM

Einar Flennstad-Jensen: EFJ

Georg Pflugbeil: GP

Gunløg Joakimsson: GJ

Hanna Hagen Bjørgaas: HHB

Hanne Hegre Grundt: HHG

Hilde Ludt: HL

Ivar Holtan: IH

Jevpeni Jakovlev: JJ

Klaus Høiland: KH

Liudmila Sergienko: LS

Leif Bond: LB

Magnus Nakkim: MN

Marit Bjørgaas: MB

Marlene Palm: MP

Marte Holten Jørgensen: MHJ

Metter Ursin: MU

Michal Ronikier: MR

Peter Kuperus: PK

Sachico Nishida: SN

Tatiana Filimonova: TF

Wilma Malik: WM

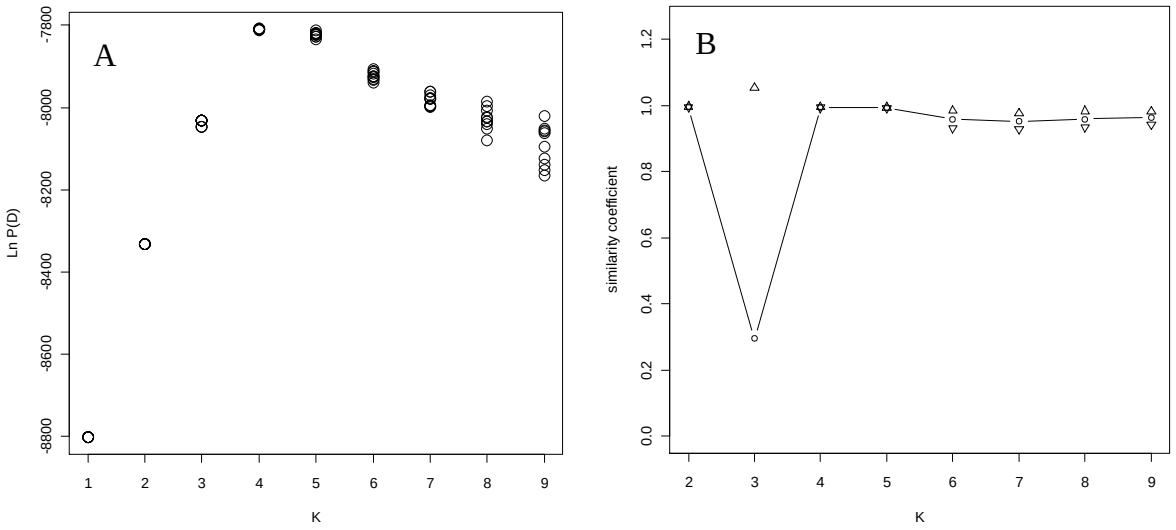

Figure S1. A) Logarithmic probability Ln P(D) of data for each K value as inferred by Structure analysis of 144 AFLP markers from *Parnassia palustris*. Each K was analysed with ten replicates. B) Plot of estimated similarity coefficients from the same Structure analyses.

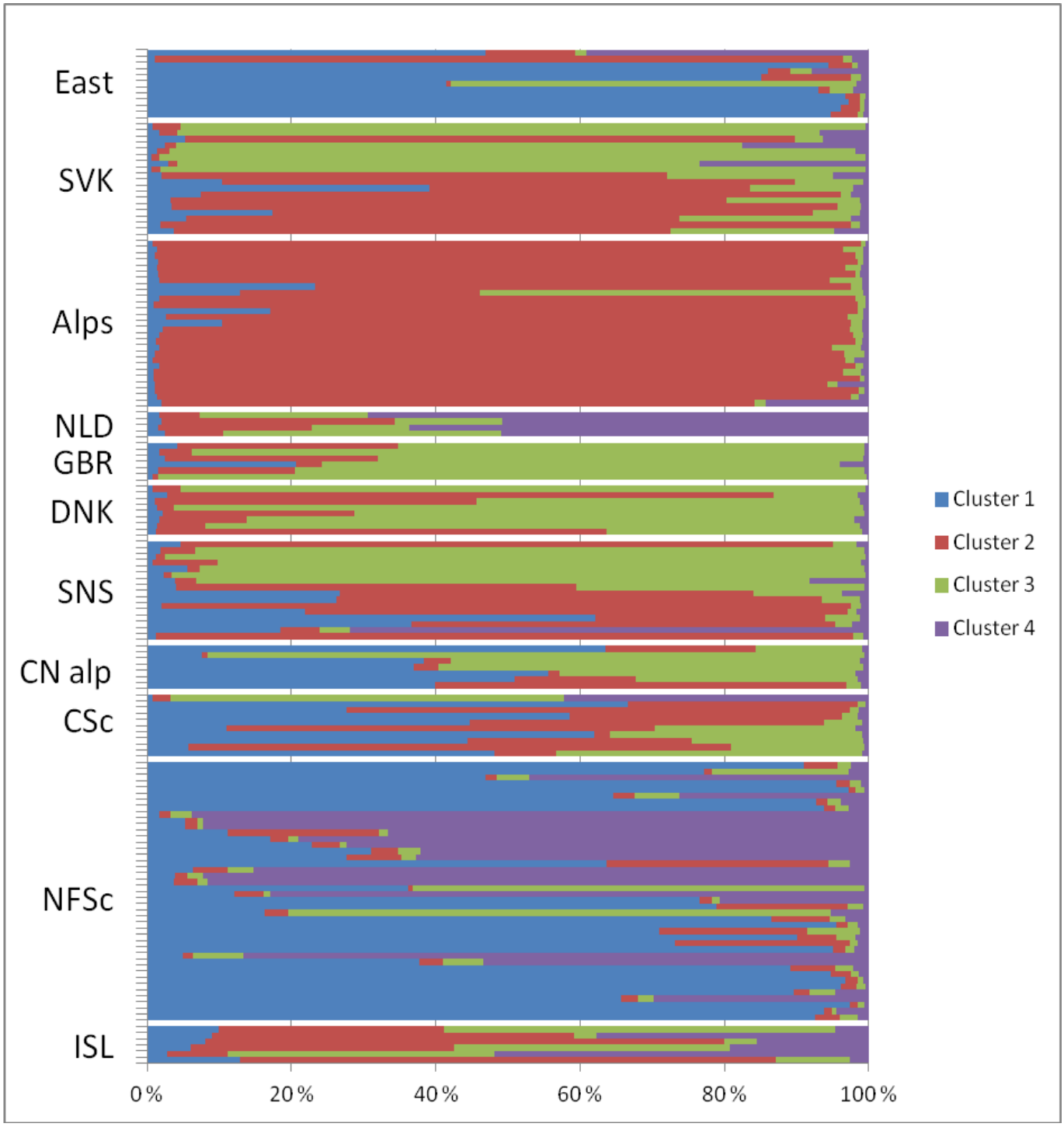

Figure S2. The allocation of European *Parnassia palustris* individuals to Cluster 1 - 4, as defined in the Structure allocation analysis based on 144 AFLP markers. Individuals are assigned to the following regions, defined in Table 3: Iceland (ISL), Northern Fennoscandia (NFSc), Central Scandinavia (CSc), Central Norway alpine areas (CN alp), Southern Norway and Sweden (SNS), Denmark (DNK), Great Britain (GBR), the Netherlands (NLD), the Alps (Alps), Slovakia (SVK), and Western Russia and Balticum (East).

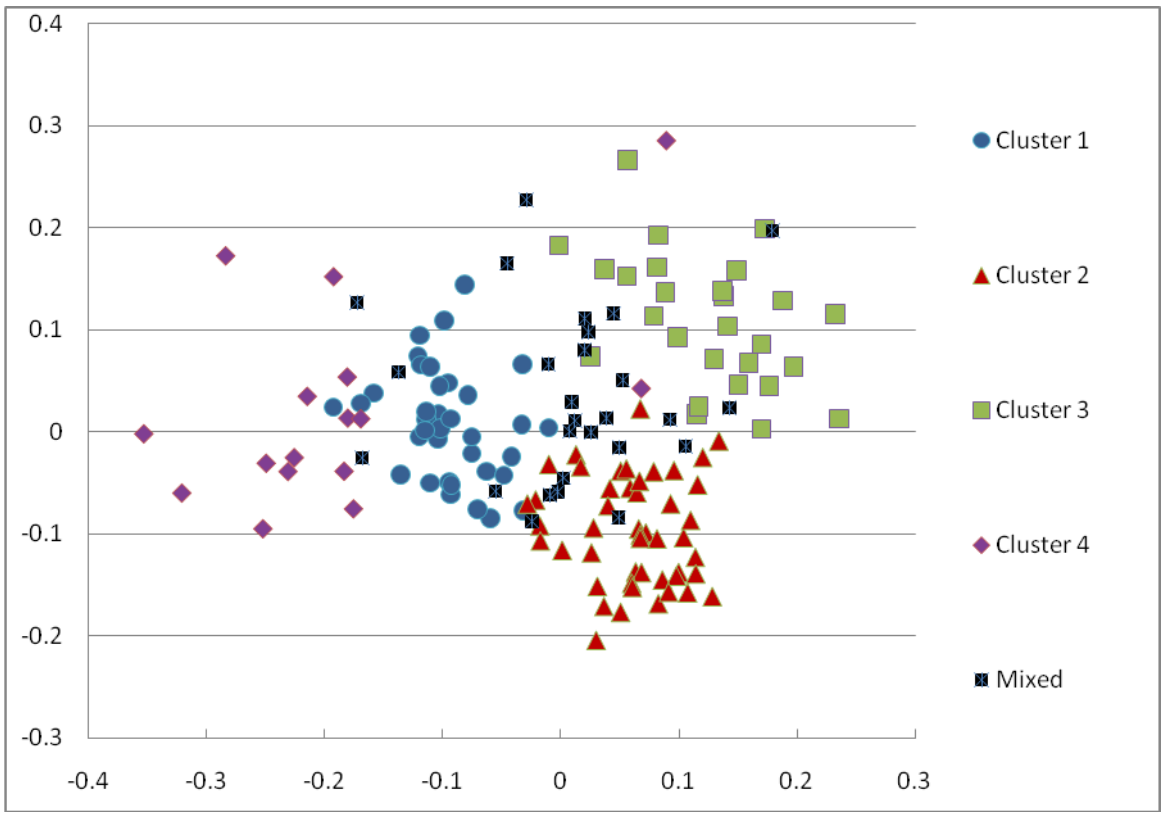

Figure S3. PCO analyses of European samples of *Parnassia palustris*, based on 144 AFLP markers using Dice as a similarity measure. Individuals are placed in one of the four Structure-defined groups if they are assigned to this group in more than 60% of the runs. Individuals that could not be placed in one of the groups are characterized as mixed.

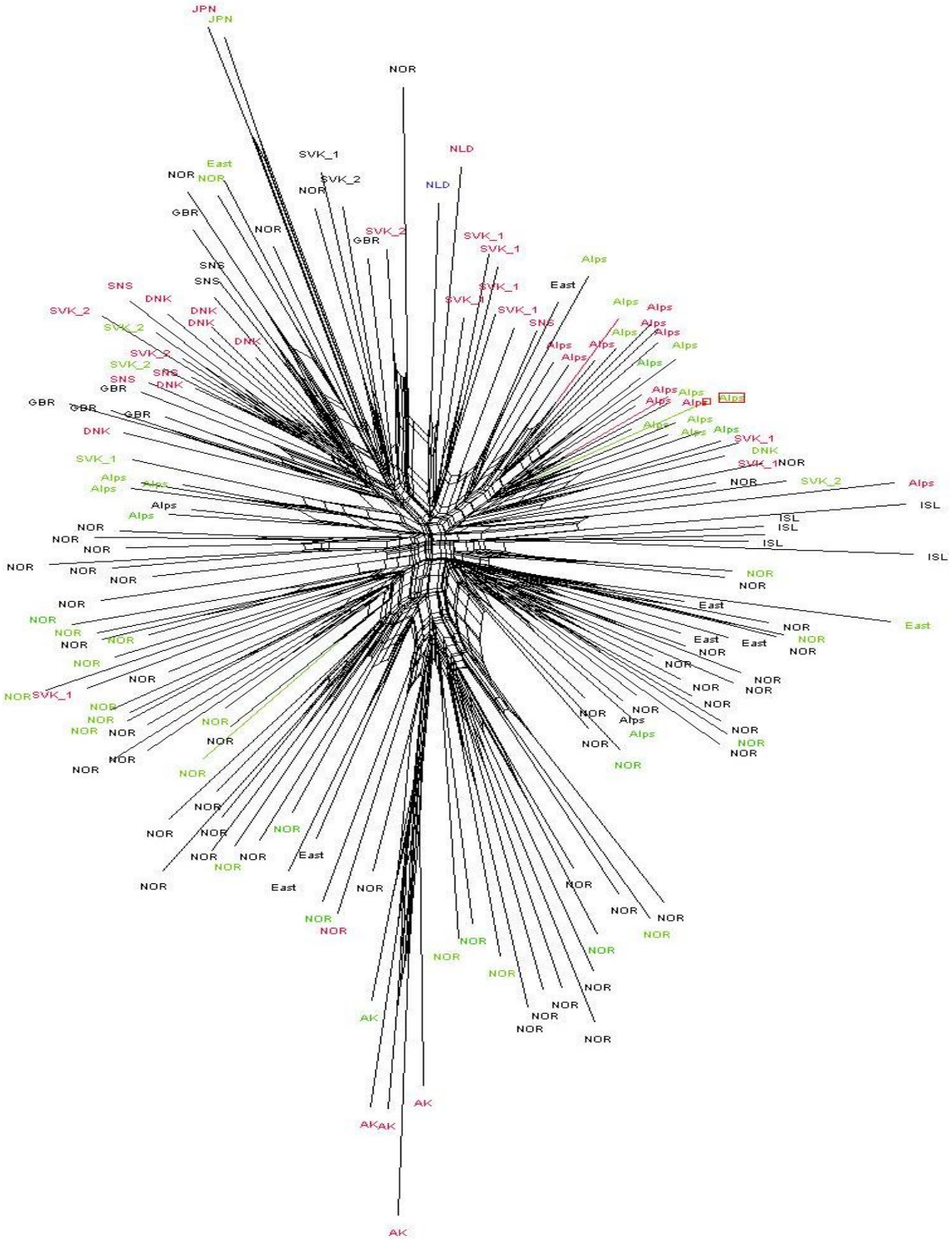

Figure S4. Neighbor-joining network of all sampled individuals of *Parnassia palustris* based on 144 AFLP markers, using uncorrected P as a measure of genetic distance. Individuals are assigned to geographical regions as defined in Table 3, except Fennoscandian individuals from localities north of 61 °N , that all are named "NOR". Ploidal level is indicated with the following colors: Red – diploid, black – tetraploid, blue – pentaploid, green – unknown.

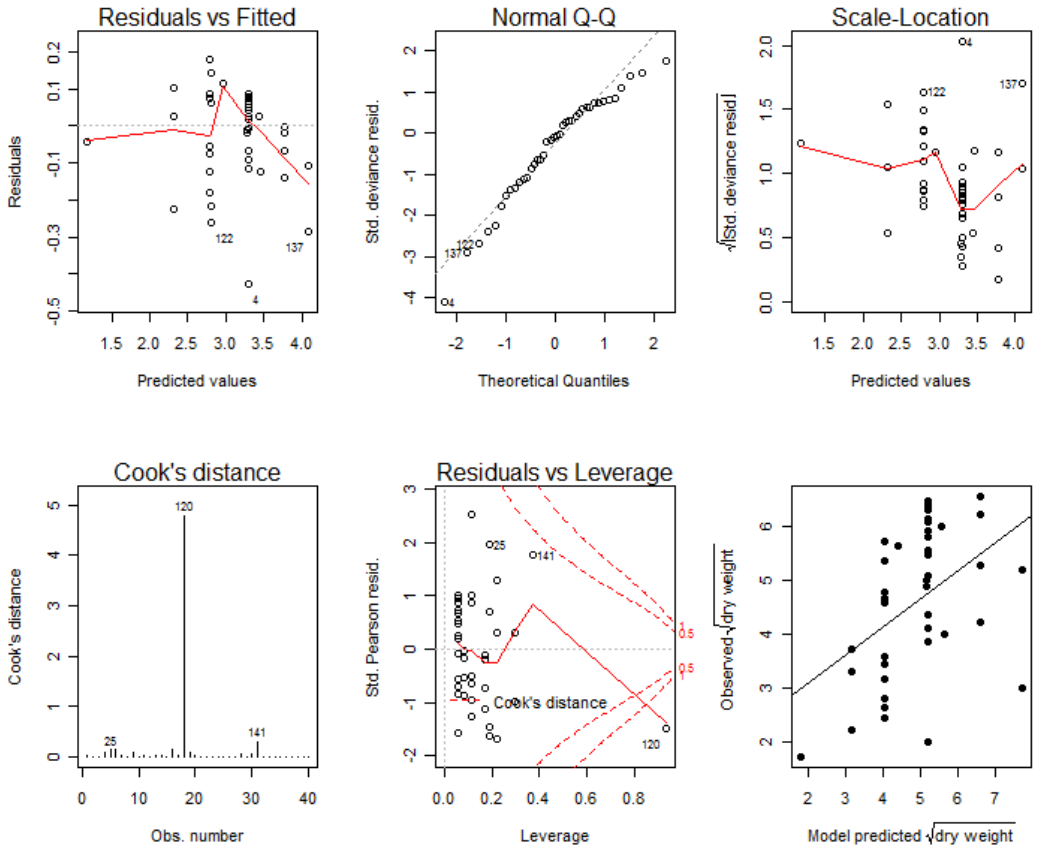

Figure S5. Residual diagnostics for the most informative and parsimonious Generalised Linear Model employed to predict dry weight of the 42 surviving plants of *Parnassia palustris* at the end of the 109 days long Experiment 1. It was assumed that the errors were following an inverse Gaussian distribution, and that a logarithmic link between the predictor and response variable was appropriate.

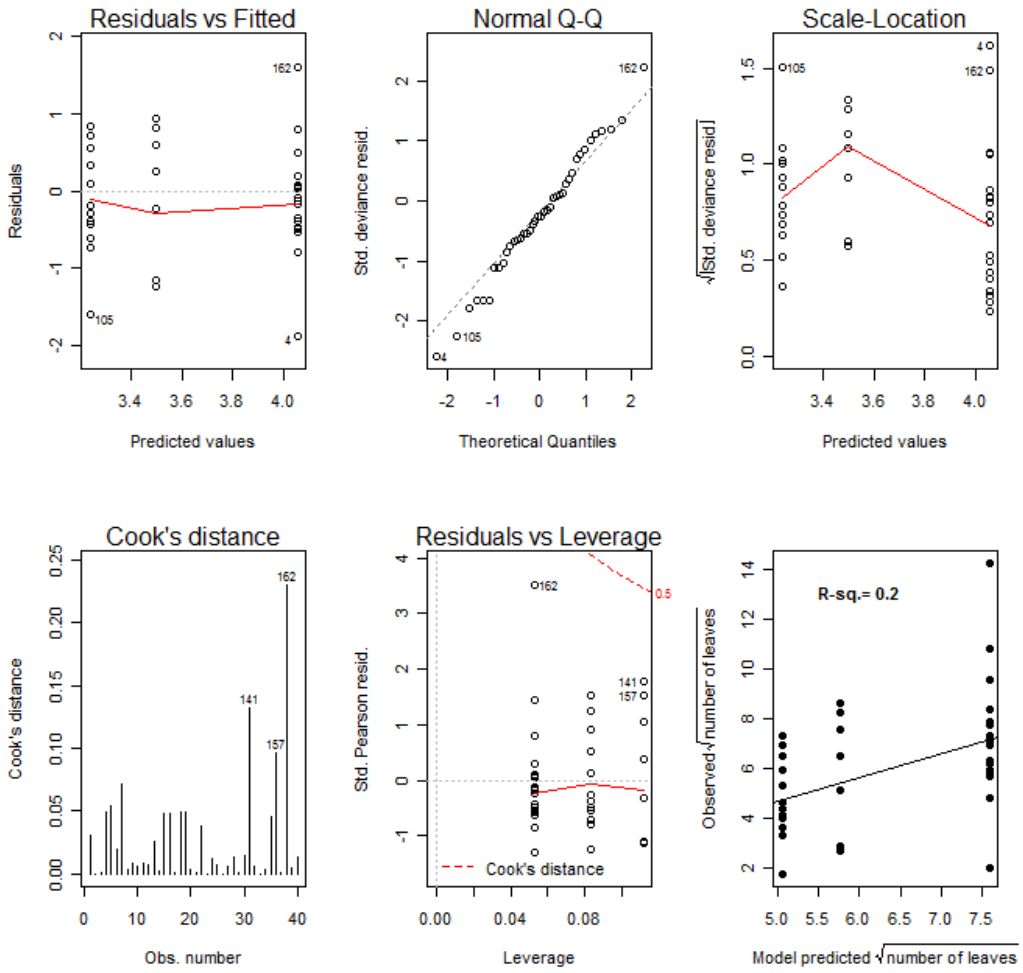

Figure S6. Residual diagnostics for the most informative and parsimonious Generalised Linear Model employed to predict number of leaves of the 42 surviving plants of *Parnassia palustris* at the end of the 109 days long Experiment 1. It was assumed that the errors were following a Gamma distribution, and that a logarithmic link between the predictor and response variable was appropriate.

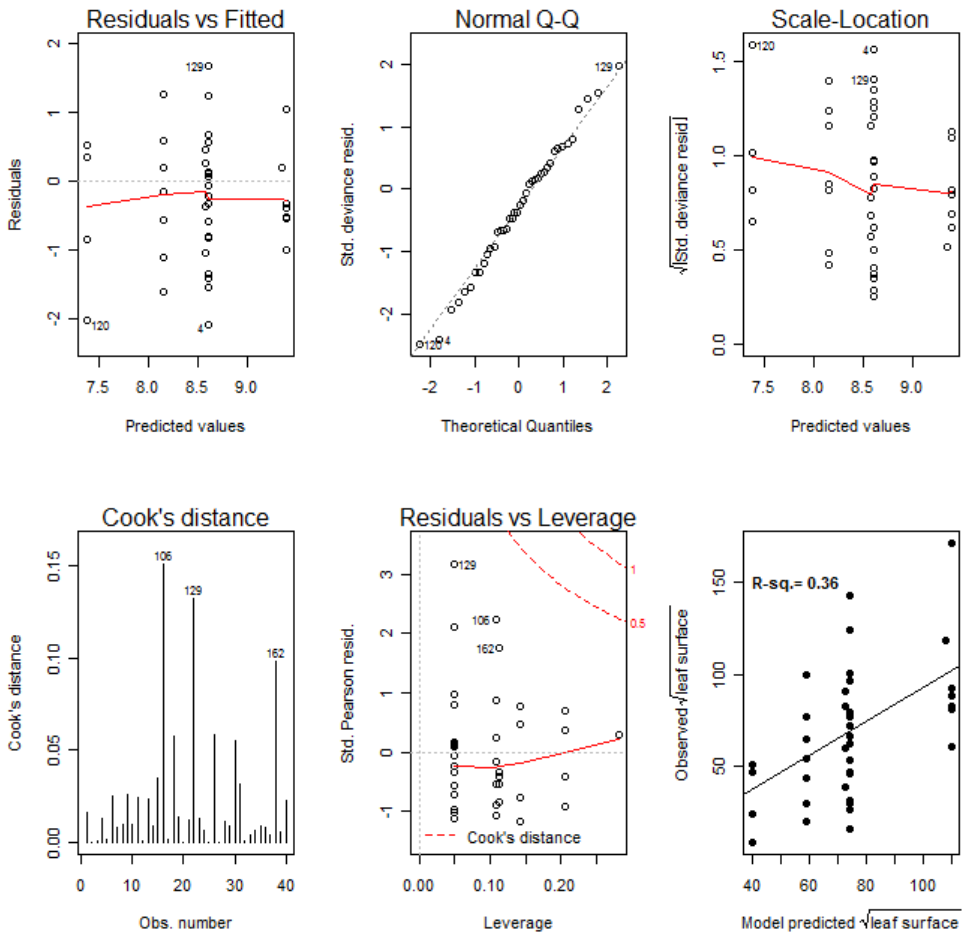

Figure S7. Residual diagnostics for the most informative and parsimonious Generalised Linear Model employed to predict leaf surface on the 42 surviving plants of *Parnassia palustris* at the end of the 109 days long Experiment 1. A logarithmic link and Gamma distribution of errors was assumed.

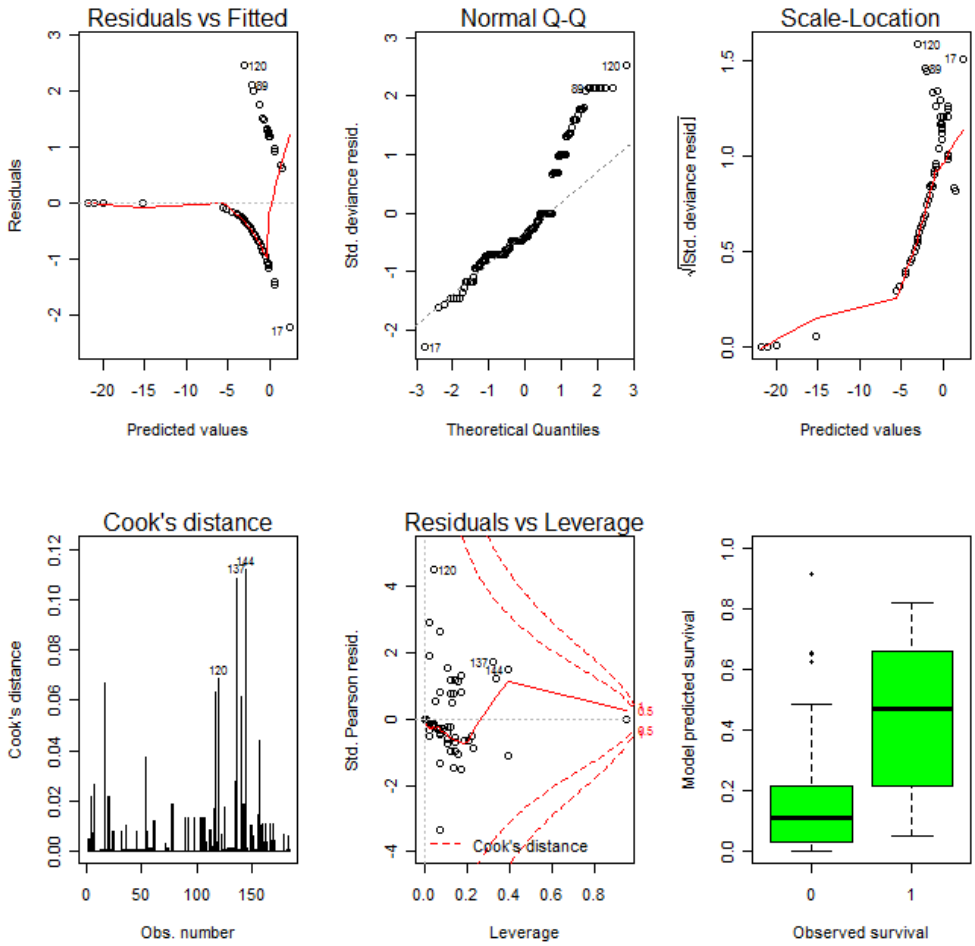

Figure S8. Residual diagnostics for the most informative and parsimonious Generalised Linear Model employed to predict survival of 184 *Parnassia palustris* plants during the 109 days long Experiment 1. A logit link and binominal distribution of errors were assumed.

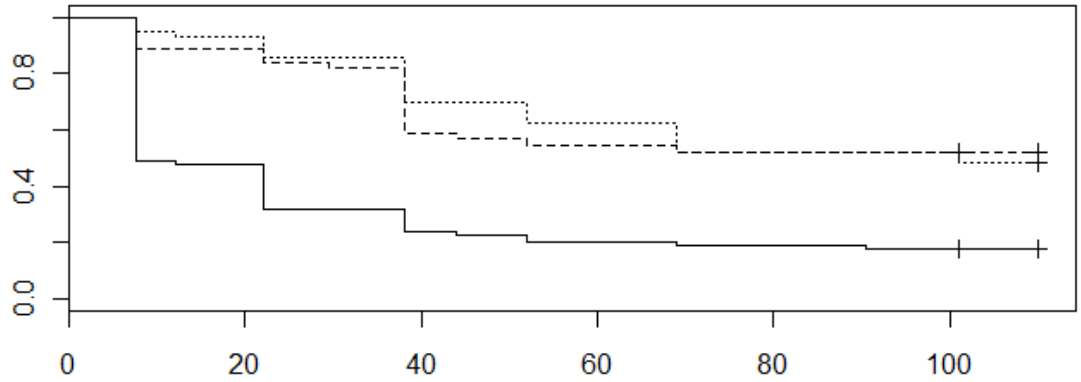

Figure S9. Kaplan-Meier curves displaying the probability of being alive after a number of days (x-axis), calculated for 222 *Parnassia palustris* plants included in Experiment 1 and 2. Full line - plants that were characterized as small in the beginning of the experiment. Medium dotted line - medium sized plants. Most dotted line - large sized plants.

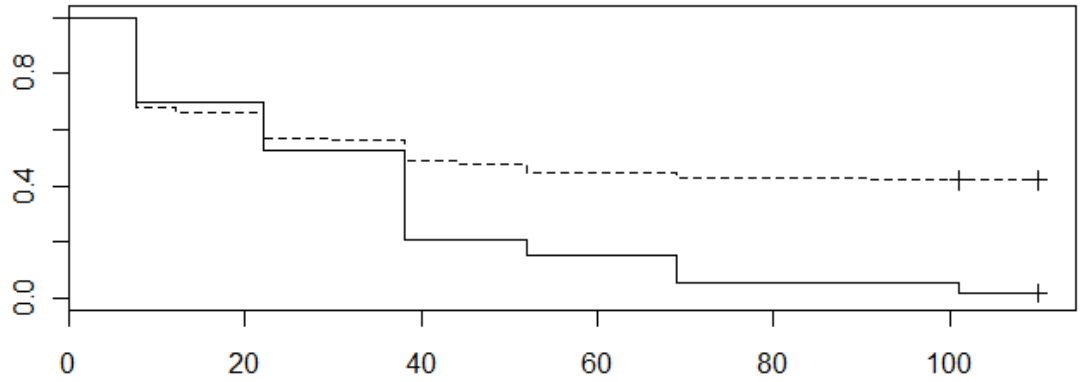

Figure S10. Kaplan-Meier curves displaying the probability of being alive after a number of days (x-axis), calculated for 222 *Parnassia palustris* plants included in Experiment 1 and 2. The dotted line represents the survival function for plants treated with the long day treatment, and the full line represents the survival function for plants treated with short day length.

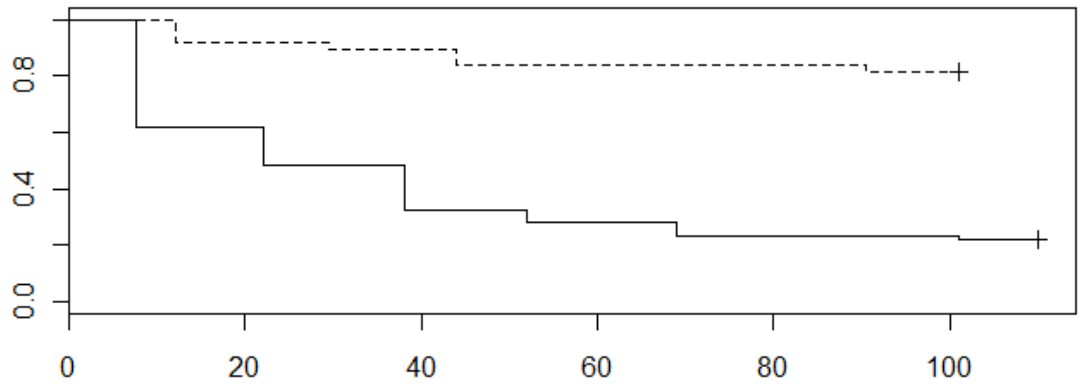

Figure S11. Kaplan-Meier curves displaying the probability of being alive after a number of days (x-axis), calculated for 222 *Parnassia palustris* plants included in Experiment 1 and 2. The dotted line represents the survival function for plants included in Experiment 2, and the full line represents the hazard rate for plants included in Experiment 1.

Table S3. The most robust predictors for various measures of growth for di- and tetraploid *Parnassia palustris* were estimated using Generalised Linear Models. The full model plus the three models with the lowest BIC values for each response variable, coefficients (Coef.) and p-value for each significant parameters are listed. The model with the lowest BIC value, i.e. the most parsimonious and informative model, is marked in green. Parameters abbreviations are: Population (Pop), Ploidy (Plo), Germination date (Ger) and Temperature (Temp). Error distribution family and link function used in the GLM is specified for each response variable. Residual deviation (Res. Dev.), Degrees of freedom (df) and the probability that the residual deviation was equal to or greater than the chi square distribution (p-value model) were calculated.

Response variable	Predictors	BIC	Parameter	Coef.	P-value	Sign.	Res. dev.	df	p-value model
Leaf number									
	$(Pop + Ger + Size + Temp)^2$ - (Ger:size +								
(Gamma, log)	Pop:Size)	407.2							
	Ger + Size + Temp	382.87							
	Ger + Temp	380.54							
	Ger	379.5	Ger1	3.4999	2E-16	**	22.274	37	0.7627
			Ger2	0.5542	0.0706				
			Ger3	-0.2612	0.4258				
Leaf surface									
	$(Pop + Ger + Size + Temp)^2$ - (Ger:size +								
(Gamma, log)	Pop:Size)	808.55							
	Pop + Size + Temp + Pop:Size	795.14							
	Pop + Size + Temp	792.08							
	Size + Temp	778.31					33.078	36	0.6083
			Size1	7.372	<2E-16	**			
			Size2	1.2051	0.0189	*			
			Size3	1.2412	0.0008	**			

Response variable	Predictors	BIC	Parameter	Coef.	P-value	Sign.	Res. dev.	df	p-value model
			Temp2	0.7831	0.0166	*			
Dry weight									
	(Pop + Ger + Size + Temp)^2 - (Ger:size +								
(Inverse gaussian, log)	Pop:Size)	356.2							
	Pop + Ger + Size + Temp + Ger:Size	346.89							
	Pop + Ger + Size + Temp	341.91							
	Ger + Size + Temp	330.69					0.68965	34	1
			(Intercept)	1.0986	1.11E-06	**			
			Germ.date2	0.5639	0.01938	*			
			Germ.date3	1.2947	1.17E-05	**			
			Size2	2.2626	4.88E-05	**			
			Size3	1.6682	4.04E-07	**			
			Temp2	0.4025	0.03797	*			
	$(Pop + Ger + Size + Temp)^3$ - (Ger:size +								
Survival	Pop:Size)	329.78							
	$(Ploidy+ Ger + Size + Temp)^3$ - (Ger:size +								
(Binominal, logit)	Pop:Size)	285.99							
	Size + Temp	185.82							
	Ger	193							
								18	
	Size	183.39					167.74	1	0.751511
			Size1	-2.1768	7.78E-12	**			
			Size2	0.8418	0.157				

Response variable	Predictors	BIC	Parameter	Coef.	P-value	Sign.	Res. dev.	df	p-value model
			Size3	2.1768	2.50E-07	**			

* - Significant at a 0.05 level
** - Significant at a 0.01 level

Table S4. Bayesian Information Criterion (BIC) values of Cox' regression models. The full models, and the most parsimonious and informative model, according to the BIC, are listed. Parameters included in the full models are experiment number (Exp), day length (Day), ploidy (Plo), germination date (Ger), size and temperature (Temp), and the response variable is the hazard rate, i.e. negative coefficients means a lowered hazard rate. Ploidy and population were highly correlated and were analysed separately. R squared (R sq) was calculated for the model with the lowest BIC. Coefficients and p-value for the parameters included in this best model is listed.

	Parameters	BIC	R sq	Parameters	Coefficients	p - value	Sign.
Ploidy							
Full model	(Exp + Day +Plo +Ger+Size +Temp)^2 – (Exp :Size +Exp :Temp + Pop: Size + Ger:Size)						
Best model	Exp + Day + Plo + Size + Temp + Plo :Temp	1403.209	0.414				
					Experiment 2	-1.953	1.30E-06 **
				Day Length 2	-0.696	1.20E-04	**
				Ploidy 4	-0.276	2.80E-01	
				Size 2	-0.895	5.00E-04	**
				Size 3	-1.538	1.00E-11	**
				Temperature 2	-0.263	2.20E-01	
				Plo 4:Temp 2	0.852	1.30E-02	*
Population							
Full model	(Day +Pop +Ger+Size +Temp)^2 – (Exp :Size +Exp:Temp + Pop: Size + Ger:Size)						
Best model	Exp + Day + Size	1394.279	0.395				
				Experiment 2	-1.976	8.10E-07	**
				Day Length 2	-0.671	2.00E-04	**
				Size 2	-0.829	7.60E-04	**
				Size 3	-1.472	1.70E-11	**

* - Significant at a 0.05 level

** - Significant at a 0.01 level

Table S5. The probabilities of proportional hazard for the parameters included in the best survival model. Parameters included in the full models are experiment number (Exp), day length (Day), ploidy (Plo), germination date (Ger), size and temperature (Temp). Ploidy and Population were highly correlated and were analysed separately.

	Parameters	**p-value**	**Sign.**
Ploidy	Exp 2	5.66E-02	
	Day 2	1.32E-06	**
	Plo 4	6.97E-01	
	Size 2	1.39E-01	
	Size 3	7.73E-03	**
	Temp 2	3.49E-01	
	Plo 4:Temp 2	8.92E-01	
Population	Experiment 2	3.40E-01	
	Day Length 2	2.50E-05	**
	Size 2	2.00E-01	
	Size 3	1.80E-02	*

* - Significant at a 0.05 level
** - Significant at a 0.01 level